

Second Edition

Exploring the X-ray Universe

Frederick D. Seward and Philip A. Charles

CAMBRIDGE

CAMBRIDGE

www.cambridge.org/9780521884839

This page intentionally left blank

Exploring the X-ray Universe

Second Edition

Capturing the excitement and accomplishments of X-ray astronomy, this second edition now includes a broader range of astronomical phenomena and dramatic new results from the most powerful X-ray telescopes.

Covering all areas of astronomical research, ranging from the smallest to the largest objects, from neutron stars to clusters of galaxies, this textbook is ideal for undergraduate students and will appeal to amateur astronomers and scientists with a general interest in astronomy. Each chapter starts with the basic aspects of the topic, explores the history of discoveries, and examines in detail modern observations and their significance. This new edition has been updated with 15 years of observations from recent space-based instruments, including results from the Hubble Space Telescope, ASCA, RossiXTE, BeppoSAX, Chandra, XMM-Newton and Swift. New chapters cover X-ray emission processes, the Interstellar Medium, the Solar System, and

gamma-ray bursts. The text is supported by more than 300 figures, with tables listing the properties of the sources, and more specialised technical points separated in boxes.

Frederick D. Seward has been leader of the High Altitude Physics Group at the Lawrence Livermore Laboratory, Director of the Einstein Guest Observer Program at the Harvard-Smithsonian Center for Astrophysics, Head of Chandra User Support and an Assistant Director of the Chandra Observatory before retiring in 2004.

Philip A. Charles is Director of the South African Astronomical Observatory, which operates SALT, the largest telescope in the Southern Hemisphere. He has recently been Head of the Astronomy Group at the University of Southampton's School of Physics and Astronomy.

Exploring the X-ray Universe

Second Edition

Frederick D. Seward

Harvard-Smithsonian Center for Astrophysics

Philip A. Charles

South African Astronomical Observatory



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org

Information on this title: www.cambridge.org/9780521884839

© Frederick D. Seward and Philip A. Charles 2010

This publication is in copyright. Subject to statutory exception and to the provision of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published in print format 2010

ISBN-13 978-0-521-88483-9 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of urls for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Contents

<i>List of acronyms</i>	<i>page</i> x
<i>Foreword</i>	xiii
<hr/>	
Chapter 1 Birth and childhood of X-ray astronomy	1
1.1 The discovery of X-rays	1
1.2 Properties of X-rays	2
1.3 The difficulties of observing X-rays from stars	2
1.4 Electromagnetic radiation and the atmosphere	3
1.5 The environment in space	4
1.6 The early years (1946–1962)	5
1.7 Sco X-1	5
1.8 An early history of the X-ray sky	6
References	10
<hr/>	
Chapter 2 X-ray emission and interaction with matter	12
2.1 Astrophysical mechanisms for generating X-rays	12
2.2 Interaction of X-rays with matter	16
References	17
<hr/>	
Chapter 3 Tools and techniques	18
3.1 X-ray detectors	18
3.2 Location of cosmic X-ray sources	24
3.3 Spectroscopy	27
3.4 Timing	29
3.5 Significant missions	30
References	37
<hr/>	
Chapter 4 Solar system X-rays	39
4.1 The production of planetary X-rays	39
4.2 Earth	39
4.3 The Moon	42
4.4 Comets	44
4.5 Jupiter	45
4.6 Saturn	47
4.7 Venus and Mars	47
References	49
<hr/>	
Chapter 5 X-ray absorption and scattering in the Interstellar Medium	51
5.1 The Interstellar Medium	51
5.2 Absorption of X-rays by interstellar gas	52

5.3 Shadows	54
5.4 Absorption by intergalactic gas	55
5.5 Scattering of X-rays by interstellar dust	55
References	59
Chapter 6 Active stellar coronae	60
6.1 The Sun	60
6.2 The Einstein stellar X-ray survey	65
6.3 The dynamo model and stars with very low L_x	68
6.4 Coronal emission from binary systems	69
6.5 High-resolution X-ray spectra	74
6.6 X-ray Doppler imaging	76
6.7 Flare stars	77
6.8 Young stars	79
6.9 Summary	82
References	83
Chapter 7 Early-type stars	84
7.1 O stars	84
7.2 X-rays	85
7.3 Stellar winds	86
7.4 X-rays from single stars	87
7.5 Colliding winds	89
7.6 Eta Carinae	90
7.7 O stars in the Carina Nebula	92
7.8 Superbubbles	93
7.9 The astronomer views the rest of the world	95
References	96
Chapter 8 Supernova explosions and their remnants	97
8.1 Introduction	97
8.2 The supernova explosion	99
8.3 Evolution to the explosion	99
8.4 X-rays from supernovae	103
8.5 Evolution of supernova remnants	106
8.6 Young shell-like remnants	111
8.7 Dominated by their surroundings: Older remnants	118
8.8 Summary	121
References	121
Chapter 9 Neutron stars, pulsars, pulsar wind nebulae, and more supernova remnants	123
9.1 Discovery and nature of neutron stars	123
9.2 The Crab Nebula	124
9.3 Rotation and spindown	126

9.4	The glitch	127
9.5	Pulsed radiation	129
9.6	Structure of neutron stars	130
9.7	Cooling	130
9.8	Pulsar wind nebulae	133
9.9	Central compact objects	135
9.10	Anomalous pulsars, soft-gamma repeaters and magnetars	135
9.11	Some supernova remnants containing a variety of neutron stars	136
9.12	Summary	140
	References	141

Chapter 10	Cataclysmic variable stars	143
-------------------	-----------------------------------	-----

10.1	Introduction	143
10.2	Geometry of accretion in CVs	146
10.3	Dwarf nova outbursts	148
10.4	X-rays from dwarf novae	151
10.5	How do CVs form and evolve?	153
10.6	Magnetic CVs	156
10.7	X-ray spectroscopy of CVs	162
10.8	AM CVn systems	164
10.9	Supersoft sources	164
10.10	Summary	168
	References	168

Chapter 11	X-ray binaries	171
-------------------	-----------------------	-----

11.1	Introduction	171
11.2	High-mass X-ray binaries	177
11.3	Low-mass X-ray binaries	201
11.4	Globular cluster X-ray sources	228
	References	239

Chapter 12	Black-hole X-ray binaries	242
-------------------	----------------------------------	-----

12.1	Introduction	242
12.2	Measuring masses	244
12.3	Observed properties	244
12.4	Soft X-ray transients (or X-ray novae)	246
12.5	Eclipsing black-hole binaries	249
12.6	Spectra of XRN	250
12.7	Black-hole spin	254
12.8	How to pick out black holes?	255
12.9	Concluding remarks	256
	References	256

Chapter 13 Normal and starburst galaxies	258
13.1 X-ray sources in the Milky Way	258
13.2 The Milky Way and the Local Group	259
13.3 Surveys of the Magellanic Clouds	261
13.4 M31	263
13.5 M33	265
13.6 Starburst galaxies	266
13.7 The centre of the Milky Way	268
13.8 Elliptical galaxies	270
13.9 Colliding galaxies	271
13.10 Source populations	273
References	275
 Chapter 14 Active galactic nuclei	 277
14.1 Introduction	277
14.2 The range of activity in ‘Active’ galaxies	279
14.3 The unified model of AGN: What is the central engine?	281
14.4 Evidence for the existence of supermassive black holes	283
14.5 Jets	292
14.6 Overall spectrum of an AGN	297
14.7 Towards the central engine: X-ray variability timescales	298
14.8 Outflowing winds in AGN	301
14.9 Summary	304
References	305
 Chapter 15 Clusters of galaxies	 306
15.1 Introduction	306
15.2 Properties of clusters: The optical legacy	306
15.3 Discovery of X-rays from the ICM	308
15.4 Evolution	308
15.5 Groups and clusters	309
15.6 X-ray observations of ICM temperature and morphology	310
15.7 The Coma Cluster: Rich, regular and almost relaxed	312
15.8 A2029: Rich, regular and relaxed, with cool centre and cD galaxy	313
15.9 A2142: Cold fronts	315
15.10 The Bullet Cluster (1E 0657-56): A merger	317
15.11 M87 and the Virgo Cluster	317
15.12 The Perseus Cluster: AGN, cavities and sound waves	319
15.13 The relation of mass, temperature and X-ray luminosity	320
15.14 Distant clusters	321
15.15 The Sunyaev-Zeldovich (SZ) effect	322
References	323

Chapter 16	The diffuse X-ray background	
16.1	Early observations, 1965	325
16.2	HEAO-1, 3–30 keV, 1977	326
16.3	Imaging the background	328
16.4	Extragalactic source populations and cosmic variance	331
16.5	Diffuse galactic emission	332
16.6	The galactic soft X-ray background, 0.2–2 keV	334
16.7	Summary	338
	References	338
Chapter 17	Gamma-ray bursts	340
17.1	Discovery	340
17.2	IPN and BATSE	341
17.3	Afterglow and precise location	342
17.4	Present understanding	347
17.5	Final thoughts on GRBs	351
17.6	The last word	351
	References	352
<i>Index</i>		354

The colour plates will be found between pages 312 and 313.

List of acronyms

AAT	Anglo-Australian Telescope
AAVSO	American Association of Variable Star Observers
ACIS	Advanced CCD Imaging Spectrometer
AGN	Active Galactic Nuclei
ANS	Astronomische Nederlandse Satelliet
ASCA	Advanced Satellite for Cosmology & Astrophysics
AS&E	American Science & Engineering
ASM	All-Sky Monitor
AU	Astronomical Unit
AXAF	Advanced X-ray Astrophysics Facility
AXP	Anomalous X-ray Pulsar
BATSE	Burst and Transient Source Experiment
BBSO	Big Bear Solar Observatory
CAL	Columbia Astrophysics Laboratory
CCD	Charge-Coupled Device
CCO	Central Compact Object
CD	Contact Discontinuity
CfA	Center for Astrophysics
CGRO	Compton Gamma-Ray Observatory
CMB	Cosmic Microwave Background
CT	Charge Transfer
CV	Cataclysmic Variable
CXC	Chandra X-ray Center
CXO	Chandra X-ray Observatory
DEM	Differential Emission Measure
DIM	Disc-Instability Model
DSS	Digital Sky Survey
DXS	Diffuse X-ray Spectrometer
EM	Emission Measure
EPIC	European Photon Imaging Camera
ESA	European Space Agency
EUVE	Extreme Ultraviolet Explorer
EXOSAT	European X-ray Observatory Satellite
FIR	Far Infrared
FPCS	Focal Plane Crystal Spectrometer
FS	Forward Shock
FWHM	Full-Width at Half Maximum
GGS	Global Geospace Science
GIS	Gas Imaging Spectrometer
GRB	Gamma-ray Burst
GSFC	Goddard Space Flight Center
GSPC	Gas-scintillation Proportional Counter
HDF	Hubble Deep Field
HEASARC	High Energy Astrophysics Science Archive Research Center (at NASA's GSFC)

HEAO	High Energy Astronomy Observatory
HETE	High Energy Transient Explorer
HETG	High Energy Transmission Grating
HEXTE	High Energy X-ray Timing Experiment
HMXB	High-Mass X-ray Binary
HR	Hertzsprung-Russell
HRC	High Resolution Camera
HRI	High Resolution Imager
HST	Hubble Space Telescope
IBIS	Imager on Board INTEGRAL Satellite
ICM	Intracluster Medium
IMBH	Intermediate-Mass Black Hole
INTEGRAL	International Gamma-Ray Astrophysics Laboratory
IP	Intermediate Polar
IPC	Imaging Proportional Counter
IPN	Interplanetary Network
IR	Infrared
ISAS	Institute of Space and Astronautical Science
ISM	Interstellar Medium
IUE	International Ultraviolet Explorer
JAXA	Japan Aerospace Exploration Agency
LANL	Los Alamos National Laboratory
LBA	Long Baseline Array
LBV	Luminous Blue Variable
LE	Low Energy
LETGS	Low Energy Transmission Grating Spectrometer
LHB	Local Hot Bubble
LINEAR	Lincoln Near-Earth Asteroid Research project
LLL	Lawrence Livermore Laboratory
LMC	Large Magellanic Cloud
LMMSC	Lockheed Martin Missiles and Space Company
LMXB	Low-Mass X-ray Binary
MCV	Magnetic Cataclysmic Variable
MCP	Micro Channel Plate
ME	Medium Energy
MIT	Massachusetts Institute of Technology
MG	MegaGauss
MOS	Metal Oxide Semiconductor
MPC	Monitor Proportional Counter
MPE	Max-Plank-Institut für Extraterrestrisch Physik
MSFC	Marshall Space Flight Center
NASA	National Aeronautics and Space Administration
NAOJ	National Astronomical Observatory of Japan
NGC	New General Catalog
NGST	Northrop-Grumman Space Technology
NIST	National Institute of Standards and Technology
NJIT	New Jersey Institute of Technology
NOAA	National Oceanic and Atmospheric Administration
NRL	Naval Research Laboratory

OGS	Objective Grating Spectrometer
OM	Optical Monitor
OSO	Orbiting Solar Observatory
PCA	Proportional Counter Array
PIXIE	Polar Ionospheric X-ray Imaging Experiment
PMS	Pre-Main Sequence
PSF	Point Spread Function
PSPC	Position-Sensitive Proportional Counter
PSR	Pulsar
PWN	Pulsar Wind Nebula
QPO	Quasi-Periodic Oscillation
QSO	Quasistellar Object
RASS	ROSAT All-Sky Survey
RGS	Reflection Grating Spectrometer
ROSAT	ROentgen SATellit
R-T	Rayleigh-Taylor
RXTE	Rossi X-ray Timing Explorer
SAAO	South African Astronomical Observatory
SAS	Small Astronomy Satellite
SAO	Smithsonian Astrophysical Observatory
SGR	Soft Gamma-ray Repeater
SIS	Solid-state Imaging Spectrometer
SMC	Small Magellanic Cloud
SN	Supernova
SNR	Supernova Remnant
SSL	Space Sciences Laboratory
SSS	Solid-State Spectrometer
SSS	Supersoft Source
STFC	Science & Technology Facilities Council
STScI	Space Telescope Science Institute
SWPC	Space Weather Prediction Center
TOO	Target of opportunity
ULX	Ultra-Luminous X-ray source
UV	Ultraviolet
UT	Universal Time
VLA	Very Large Array
WFC	Wide Field Camera
WC	Wolf-Rayet (carbon-rich)
WR	Wolf-Rayet
XLF	X-ray Luminosity function
XMM	X-ray Multi-Mirror Mission
XRO	X-ray Outburst
XRB	X-ray Binary
XRT	X-ray Telescope
ZAMS	Zero Age Main Sequence

Foreword

This is a book about X-ray astronomy. We take a historical perspective because this is how we saw it happen and because this gives a feeling for the observable universe. In a table listing all members of a class of objects, the brightest source does not stand out, but in the first observation, it is a splendid object and remembered fondly by those involved in the discovery.

Some 50 years ago X-rays from stars other than our Sun were unknown and unexpected by all but a few pioneering scientists. Since the discovery of cosmic X-rays in 1962 the field has grown at an astonishing rate. Since the first edition of this book, published in 1995 and including results from the first X-ray telescopes, the sensitivity of X-ray observations has increased dramatically. In 1999 the Chandra and XMM X-ray observatories were launched and, in 10 years of operation, have produced X-ray images of comparable angular resolution to those obtained by the largest ground-based observatories. More importantly, X-ray spectroscopy of sufficient resolution to allow comparison with spectra at other wavelengths has become possible. Technical improvements in dispersive spectroscopy mean that high resolution X-ray spectra of faint sources have become available for the first time. This has helped propel X-ray astronomy to its rightful place as a sub-discipline of astronomy, where a knowledge of truly multiwavelength results is necessary for the study of any class of objects. This book, however, is about X-rays. Observations at other wavelengths are sometimes included but space limitations mean that thorough coverage is beyond the scope of the book, although we try to give suitable references and links.

We are observers, so, although this book is primarily about the science of X-ray sources, there are excursions into instrumentation and techniques. Our research has been mostly galactic, as all the bright X-ray sources are nearby and have been available for study from the earliest days of this field. This means that there is somewhat less detail in chapters concerning sources at much greater

distances, but our understanding of the early universe and cosmology is developing rapidly at the present time. We have each written chapters concerning our respective special interests but could not cover all aspects. We were trained in nuclear physics (FDS) and optical/X-ray astronomy (PAC). These backgrounds probably shine through in some of the material covered. We hope all major areas touched by X-ray astronomy are included.

The material is written for someone with an astronomical background, but each chapter starts simply so those with no background can understand the basics. For people who do not wish to read all the details, we have separated some technical discussions into boxes that accompany the text. An acquaintance with physics and mathematics is helpful in understanding some of the boxed material but the main text is presented so that it can be followed without the boxes.

Most of the illustrations have already appeared in the scientific literature. A few are from our own work and we thank our colleagues who have generously given material generated by their research. Although many references have been added, we have not given extensive accreditation in order to make the text more readable (and not outrageously long). The bibliography for each chapter includes some popular articles and scientific reviews. Material obtained from the Internet is referenced by the Internet address (URL) and sometimes by a more-descriptive publication.

Some scientists who have made key observations or developed key theories are mentioned in the text but many are not. We apologise to our friends and colleagues for the omissions. We also recognise that astronomers only contribute part of the effort necessary for a successful space astronomy mission. There is planning, funding, building, calibration, launch, control during observations, data transmission and collection, and the data must be made available to observers in an understandable format. The scientific

process also includes editors, referees, and those who evaluate proposals. X-ray astronomy now is very much a group effort, a national or international effort for the large missions. Our colleagues at NASA and ESA are career scientists who manage ongoing programs, administer the funding, and plan the future of X-ray astronomy. This is a difficult and, at times, a thankless job. Successful proposers think winning is their just due. Unsuccessful proposers think the process is flawed. It is easier to make enemies than friends. We wish them adequate funding and success with future projects.

We would like to thank those who took the trouble to contact us with comments about and corrections to our 1st edition, for which we are very grateful. We also thank our many colleagues and friends for providing us with material, particularly diagrams, that we have included in this new edition. In particular, we appreciate helpful discussions with Jeremy Drake, Bill

Forman, Dave Huenemoerder, Dong-Woo Kim, Andrea Prestwich, John Raymond, Pat Slane, Jan Vrtilek, and Scott Wolk. We hope their input has been fairly treated.

We, the authors, have been observers of celestial X-rays since 1965 (FDS) and 1972 (PAC). We miss the early days when a person could easily name from memory all the known X-ray sources and when the nature of many were unknown. There are now so many instruments and sources that we are challenged to even list the observatories or categories of objects. Many of the major discoveries have been serendipitous in nature, leading to an exciting, but at times chaotic, growth in the field (a process reflected perhaps across all of observational science). We have nevertheless organised the material and note with satisfaction that the underlying physical mechanism of most cosmic sources is now, at some level, understood. We hope that you gain knowledge and enjoyment from this book.

Phil Charles – South African Astronomical Observatory
Fred Seward – Smithsonian Astrophysical Observatory

Chapter I

Birth and childhood of X-ray astronomy

I.1 | The discovery of X-rays

On the second story of the building at Röntenring 8 in Würzburg, Germany, there is a plaque: ‘In diesem Hause entdeckte W. C. Röntgen im Jahre 1895 die nach ihm benannten Strahlen’ – In this building, in the year 1895, W. C. Röntgen discovered the radiation named for him. Here was the laboratory of Wilhelm C. Röntgen, a 50-year-old professor of physics, who was studying phenomena associated with electrical discharge in gasses. On the afternoon of 8 November, working alone in his laboratory, he noticed a curious phenomenon. When high voltage was applied to the electrodes in the partially evacuated glass discharge tube, he noticed a faint glow from a fluorescent screen placed at the other end of the laboratory table. The room was dark and he had previously covered the tube with black cardboard so no light would escape. Why was the screen glowing?

That evening he verified that the discharge tube was indeed the source of the energy that caused the screen to glow, and that no visible radiation was escaping from the shrouded tube. He quickly found that the unknown radiation would pass through paper, wood, and aluminum but was stopped by heavy metals. Then, when holding a lead disc in front of the screen to observe its shadow, Röntgen also saw the shadow of bones in his hand! In a week he had measured the basic characteristics of this new form of radiation. He persuaded his wife, Bertha, to hold her hand

steadily over a photographic plate for 15 minutes, making a picture showing hand and finger bones as grey shadows and the shadow of her ring sharp and black. He sent this picture and a description of his results to be published and to other scientists. He called these new rays ‘X-rays’ but others called them ‘Röntgenstrahlung’ – Röntgen radiation. The medical applications were immediately obvious and commercial X-ray machines were soon available.

The discovery of X-rays is one of the most famous serendipitous discoveries of science (Colour Plate 4). If Röntgen had been funded to investigate ways to help doctors in hospitals set broken bones, it is unlikely that he would have pursued this line of research.

In 1901, Röntgen was awarded the first Nobel Prize in physics. In 1990, a German satellite devoted to X-ray astronomy was placed in orbit and operated for 10 years. The satellite was named ROSAT (for Röntgen Satellite) and was taking data on the 100th anniversary of Röntgen’s discovery. In 2002, Riccardo Giacconi was awarded the Nobel Prize in physics for pioneering work in X-ray astronomy. The authors of this book have had the pleasure of knowing Riccardo and, years ago, of observing the source he discovered, Sco X-1. Now, 115 years after Röntgen’s discovery, we are finishing the initial period of exploration, and X-ray astronomy has become part of the general field of astronomy. Like traditional optical astronomers, we work with telescopes that record hundreds of sources in each field. It is still fun.

1.2 | Properties of X-rays

We know now that X-rays are a form of electromagnetic radiation, like visible light, but the individual quanta of radiation, the photons, have energies a thousand times that of optical photons. Visible light, in general, does not penetrate matter. The photons are scattered from or absorbed at the surface of opaque objects. Visible light will also pass through transparent substances and will reflect from smooth surfaces. X-rays, on the other hand, go right through the surfaces of all substances. The individual photons, traveling in straight lines, either interact with individual atoms or pass through unaffected. The probability of interaction increases with Z , the atomic number of the element. Thus, in Röntgen's picture of Bertha's hand, the X-rays passed easily through the carbon ($Z = 6$) and oxygen ($Z = 8$) of flesh, but many photons were absorbed by the calcium ($Z = 20$) of bones, and all were absorbed by the gold ($Z = 79$) ring that was on her finger. The probability of interaction also depends on photon energy. More energetic photons are less likely to be absorbed.

X-rays are usually generated by accelerating an electron beam with high voltage and directing the beam to strike a tungsten target. The energy of the X-ray photons is measured in kilo-electron volts (keV), the voltage used to accelerate the electrons. A potential of 100 000 volts, for example, is capable of producing 100 keV X-rays.

1.3 | The difficulties of observing X-rays from stars

Röntgen's X-rays had energies of 30–50 keV, about the same energy as X-rays used today for medical diagnostics. Astronomical X-rays are much less energetic and more easily absorbed. Most X-rays from cosmic sources cannot penetrate even the thin outer layers of the Earth's atmosphere. It is thus impossible to observe X-rays from astronomical sources with ground-based instruments. Even from mountain tops, airplanes, and simple balloons, attempted observations are hopeless.

Hence, observing from above the atmosphere is essential in this field. To see any X-rays at all, it is necessary to be above 99 per cent of the atmosphere, and to detect X-rays in the band where sources are most prominent, all but one millionth of the atmosphere must be below the instrument.

Cosmic X-ray sources are most clearly detected in the range of 0.5–5 keV in photon energy (or wavelength of 25–2.5 Å). By earthly standards, these X-rays are 'soft' and easily stopped by a small amount of material. For example, three sheets of paper or 10 cm of air at one atmosphere pressure will stop 90 per cent of 3 keV X-rays. The higher the energy, however, the more penetrating, or harder, the X-rays. A rocket is needed to observe 3 keV X-rays, which cannot be seen at altitudes below 80 km, whereas 30 keV photons will penetrate to 35 km altitude, which can be reached by the highest-flying balloons. The instrument should be above 200 km to observe X-rays with energies below 1 keV in a direction parallel to the Earth's surface, as would be desirable in a survey. A few cosmic sources emit hard X-rays and have been observed with balloons but almost all work is now done with satellite-borne instruments.

It was not a trivial matter to build the first instruments that were large enough to be sensitive yet small enough to fit within a rocket or balloon payload. The instruments not only had to withstand the rigours of launch but also had to operate in a vacuum or near vacuum. Time and trial and error were needed to develop the first survey instruments. To detect, for the first time, a phenomenon that many people believed impossible, took confidence that the instruments were operating properly. Much of the early work was done by nuclear physicists, who were familiar with the type of detectors used to register the X-rays.

Because X-rays are a form of electromagnetic radiation like visible light, they can be produced by the same processes. Because the photon energies are 1000 times greater than that of optical photons, the process must be correspondingly more energetic to produce X-rays. So, if X-rays are generated in a thermal process, the temperature must be of the order of 1000 times greater than that in places where light is produced. Thus, a search for cosmic X-ray sources is a search for

material at temperatures of millions of degrees, in contrast to the familiar stars with surface temperatures of thousands of degrees. Until 1962, very few astronomers believed that the Universe contained objects capable of generating detectable amounts of high energy radiation and little was expected from the first observations.

These ideas changed dramatically in the early 1960s with the discovery that there were indeed many discrete, powerful sources of astronomical X-rays. Some produce X-rays by processes unimagined until the observations forced people to consider new kinds of cosmic objects and new methods of energy production.

1.4 Electromagnetic radiation and the atmosphere

It is not an accident that our eyes operate in the narrow waveband 4000–8000 Å. Not only is a large fraction of the energy of the Sun radiated in this band but the Earth's atmosphere is almost transparent throughout this 4000 Å-wide waveband. Figure 1.1 shows the electromagnetic spectrum from radio waves to gamma-rays and depicts the depth to which each frequency can

Table 1.1 Atmospheric transmission of X-rays.

altitude (km)	energy (keV)	transmission, source overhead	transmission, source 90° from vertical
40	20	0.15	0.00
40	30	0.64	0.00
150	1	0.98	0.64
200	0.2	0.99	0.82

penetrate the atmosphere. Over twelve decades of the spectrum, from gamma-rays to the far infrared (FIR), there is only a very narrow band of radiation that reaches the Earth's surface essentially unscathed. The 'opacity' (potential to absorb) of the atmosphere is the principal difficulty facing astronomers wishing to study radiation from the stars at wavelengths outside the visible band. Until this century, the visible part of the spectrum was all that was available for study of the heavens. Only radio astronomy was able to develop at all using ground-based instrumentation, although it is now possible to undertake infrared observations from high altitude observatories through some windows less affected by water vapor.

Figure 1.1 shows the height above sea level to which radiation of each wavelength can penetrate. All radiation from the extreme ultraviolet (UV) (at 1000 Å) to X-rays to high-energy γ -rays (at 10^{-4} Å) fails to penetrate below an altitude of ~ 30 km. It is the requirement of observing above the atmosphere that makes the study of the X-ray Universe a modern one. At 40 km altitude, typical for balloon flights, the atmospheric transmission of 30 keV X-rays from a source directly overhead will be ≈ 60 per cent. If the source is 60° from the vertical, transmission is ≈ 36 per cent. A source 90° from the vertical cannot be detected. (This direction is the 'horizon' on Earth's surface but, at high altitude, the actual horizon can be well below this direction.) Table 1.1 gives transmissions for other altitudes and energies. It shows that 20 keV is about the low energy limit of balloon-borne detectors and that a rocket has to be above ~ 150 km to perform a useful scan. At 200 km, the transmission of even the softest X-rays is high. Because

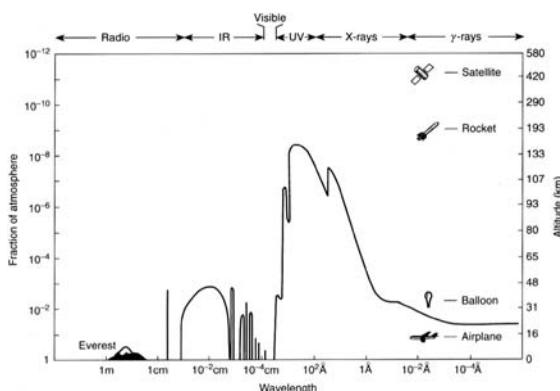


Fig 1.1 Transmission of electromagnetic radiation by the atmosphere. The solid line shows the altitude by which half the radiation from space has been attenuated. Just below this line virtually all the radiation is absorbed. Only radio, optical, and some narrow bands of infrared radiation can reach the Earth's surface. High energy γ -rays can be observed using balloons, but rockets or satellites are necessary for X-ray or UV detection.

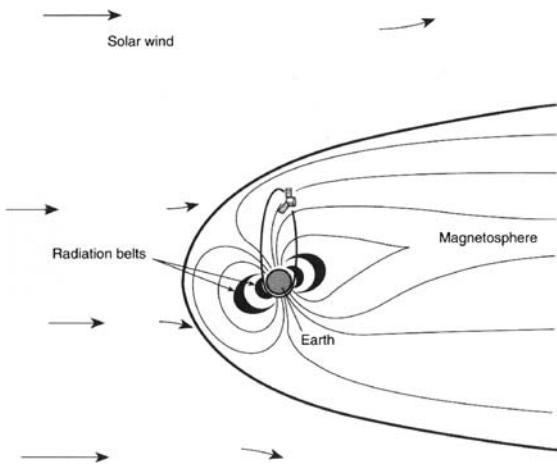


Fig 1.2 Earth and its magnetosphere. Magnetic field lines are connected to the Earth's poles and confined inside the bow shock, which is usually ~ 12 Earth radii distant in the solar direction. The inner and outer radiation belts are shaded. A spacecraft in the eccentric orbit shown spends most of its time above the intense zones of trapped particles but is always within the magnetosphere.

satellites operate at higher altitudes to avoid atmospheric drag, atmospheric attenuation does not limit satellite X-ray surveys.

1.5 | The environment in space

Two restrictions govern orbits selected for X-ray observatories: (i) although soft cosmic X-rays can be detected at altitudes above 150 km (reached easily by rocket), a satellite in a circular orbit must have an altitude greater than 400 km or the drag of the tenuous upper atmosphere will soon lead to a fiery re-entry and (ii) because X-ray detectors are sensitive to energetic charged particles, observations are best done where particle fluxes (and detector backgrounds) are at a minimum.

Figure 1.2 shows (not to scale) the space environment in the vicinity of Earth. Inside the parabolic magnetopause, space is dominated by Earth's magnetic field. Outside, energy from the Sun, in the form of solar wind, magnetic field, and energetic particles, governs conditions. A bow shock forms just outside the magnetopause. The solar side of the magnetosphere boundary is normally $10\text{--}12 R_{\oplus}$ (Earth radii) distant but can be

pushed closer by solar activity. At a lower altitude, Earth is girdled by the pitted-olive shaped Van Allen radiation belts. Here, protons and electrons trapped in the Earth's magnetic field have both high energy and long lifetimes. The best orbits for observation minimise time spent in these zones, which extend from $\sim 0.2\text{--}5 R_{\oplus}$. Indeed, detectors must be powered off so as to prevent damage when in the heart of the belts.

A near-Earth equatorial orbit with altitude ≈ 500 km is below the belts except for a region over the South Atlantic Ocean where Earth's field is weak and trapped particles dip to lower altitudes. Earth's field keeps solar particles from this equatorial region and reduces the flux of cosmic rays. In addition, the solid Earth stops half the high-energy cosmic rays that penetrate the magnetic field. The consequent low background is an advantage, but in a near-Earth orbit, only half the sky is visible at any one time. Except for targets near the poles of the orbit, this means that the observations will be repeatedly interrupted for a large part of every ≈ 100 -min orbital period. (This is exactly what happens with the Hubble Space Telescope.) Also, all detectors have to be turned off when passing through the South Atlantic Anomaly where charged particle flux is high. This causes another interruption on many orbits. Near-Earth polar orbits have also been used, usually because of launch site or data-receiving station locations. In these orbits, particle-induced backgrounds are low at low latitudes but high in polar regions due to precipitating solar and auroral particles.

A highly eccentric elliptical orbit will allow the spacecraft to spend almost all of the time well above the Van Allen zones. If beyond the magnetosphere, however, there is no protection from solar particles of all energies and background rates can be high after solar flares. Indeed, exceptionally large flares on the Sun can inject a huge number of energetic particles out into the Solar System. These affect the Earth by causing telecommunication problems and also spectacular auroral displays at high latitudes. Such a large flare can sometimes result in the need to turn spacecraft instruments off for several days, whereas within the magnetosphere, some shielding is provided from solar particle events. No orbit is free

from transient solar particle fluxes and spacecraft operators must be vigilant to avoid damage to the sensitive detectors.

1.6 | The early years (1946–1962)

The first technology useful for research above the atmosphere was that of the V2 rockets available after World War II. With these, the U.S. Naval Research Laboratory (NRL), under the direction of Herbert Friedman, was able to reveal the Sun as a powerful source of UV and X-radiation. Perversely, this discovery actually caused many scientists to lose interest in the search for other sources of X-rays, as they realised that the Sun appears as a bright source only because it is extremely close to us. A calculation of the intensity of radiation expected at the Earth from the nearest stars (assuming that they are comparable emitters of X-rays to the Sun) showed that the instrumentation available in 1960 would have had to be about a factor of 10^5 more sensitive to detect such objects. Worse still, if the stars were more distant, at a typical distance of 1 kiloparsec (kpc) or about 3000 light years, then a 1960 observation would only have been capable of discovering a process which was producing 10^{11} times the X-ray luminosity of our Sun.

Most of the rocket observations of the 1950s were devoted to more detailed studies of the Sun, although the NRL group did search (without success) for other cosmic sources. Even so, several groups kept working to develop more sensitive instruments. In the end, it was a group at American Science and Engineering (AS&E), led by Riccardo Giacconi, that was successful in the first detection of a powerful cosmic source of X-radiation.

1.7 | Sco X-1

The official purpose of the AS&E experiment was to search for X-rays from the Moon, which were expected to be produced by the energetic solar wind particles striking the lunar surface, with perhaps some fluorescence from solar X-rays. A

positive result would provide valuable information about the nature of the lunar surface; an area receiving much publicity and support at the time with America's commitment to a manned lunar landing within the decade. In addition, it was planned to scan a large region of sky in a search for non-solar sources of X-radiation. The first launch of this new instrument took place in October 1961. The rocket launch was perfect but the doors, designed to protect the X-ray detectors during launch and passage through the atmosphere, failed to open! In the early days, equipment was simple but often unreliable.

The second launch of the AS&E instrument, on a new Aerobee rocket, took place from White Sands, New Mexico, on 18 June 1962 and this time the doors functioned perfectly. Two of the three X-ray Geiger counters worked well and, although they failed to detect any X-rays from the Moon's surface, they made the first detection of a powerful cosmic X-ray source (Giacconi, Gursky & Paolini, 1962). This source subsequently became known as Sco X-1, the first-discovered source in the constellation Scorpius. As Richard Hirsch (1983, p. 46) comments in his history of X-ray astronomy, 'Observing Sco X-1 was the reward nature offered to scientists willing to gamble on a long shot'.

Interpretation of the data was not straightforward. With this detector, precipitating electrons could produce a signal similar to that of an X-ray source. By realising that the observed signal of $100 \text{ photons cm}^{-2} \text{ s}^{-1}$ was indeed caused by an extrasolar X-ray source, Giacconi and colleagues captured the interest of the astronomical community and started an exploration that has uncovered some truly remarkable objects.

The unusual nature of Sco X-1 was clear as soon as it had been roughly located. Figures 1.3 and 1.4 contrast the X-ray and optical appearance of Sco X-1. The source dominates an early rocket X-ray survey. An optical picture, however, containing Sco X-1 shows nothing unusual whatsoever. Until an accurate location of the X-ray source was obtained (Gursky *et al.*, 1966), astronomers had not a clue as to the nature of this source. Sco X-1 is an object which stands out like a beacon to a small X-ray detector but is visually four hundred times fainter than the faintest star that can be seen with the naked eye. In every square degree of

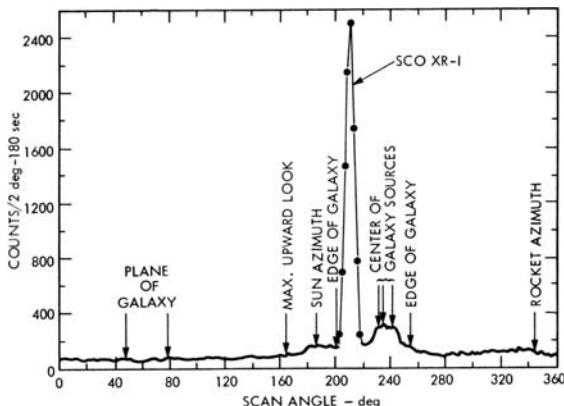


Fig 1.3 Three minutes of data from a rocket-borne X-ray detector flown in October 1967. This shows the counting rate of the detector as it scanned a great circle containing the source Sco X-1 and a cluster of sources in the direction of the galactic centre. The detector field of view was 5° by 30° . The Sun was below the horizon. The signal from Sco X-1 is very strong. (from Hill *et al.*, 1968).

the sky there are about one hundred stars visually brighter than Sco X-1.

1.8 An early history of the X-ray sky

After the discovery of Sco X-1, X-ray astronomy progressed rapidly. Evidence for two weaker sources was found on 12 October 1962 by the AS&E group (Gursky *et al.*, 1963). The NRL group confirmed and located one of these sources on 29 April 1963 using a rocket-borne detector (Bowyer *et al.*, 1964). It was identified right away as the Crab Nebula, a well-known young supernova remnant in our galaxy, and high energy X-rays from this source were detected on 21 July 1964 by George Clark (1965) of the Massachusetts Institute of Technology (MIT). (This was the first detection of high energy radiation from an extrasolar source with a balloon-borne detector.)

Astronomers were thus forced to recognise that there were many objects at stellar distances which were strong, unbelievably strong, sources of high energy photons. Small areas of the sky were then explored with great enthusiasm using rockets and balloons. The ‘big picture’ was not revealed

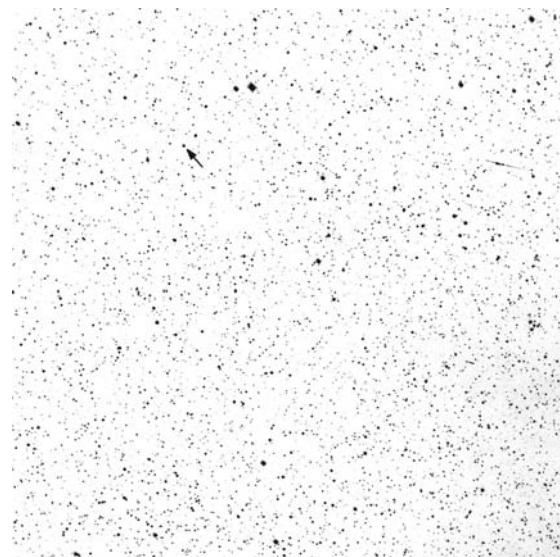


Fig 1.4 One square degree of the sky from the Palomar Sky Survey. The 13th magnitude star indicated with an arrow is the optical counterpart of Sco X-1.

until the first survey with the Uhuru satellite, launched on 12 December 1970.

The nature of the X-ray sources and the manner in which energy was generated was not obvious. It was first necessary to obtain precise locations of X-ray sources, leading to identification with optical or radio objects. The next steps were to measure the X-ray spectra and light curves to determine the emission mechanism.

Some of the brighter sources in our Galaxy radiate 10 000 times as much energy as does the Sun across all wavelengths. Almost all (99.9%) of this energy appears as X-rays. Sco X-1 is such a source. The optical counterpart is a 13th magnitude star, invisible to the naked eye and even to small telescopes. The only visual clues to its unusual nature are a blue-violet colour and an irregular variability marked by occasional rapid flickering. No optical surveys previous to the X-ray detection had indicated anything unusual. Even after the optical counterpart had been identified, the Sco X-1 system was not understood.

A convincing explanation of its nature was not found until 1971 when Uhuru discovered and measured the peculiar X-ray variation of another

source that lies in the southern sky in the constellation Centaurus. This source, Cen X-3 (or 4U1118-60), is an X-ray-bright object at a declination of -60° . Although bright enough to be detected by a rocket-borne instrument, it was below the horizon for sounding rockets launched from the main U.S. facility at White Sands. It was clearly accessible, however, to those using launchers in Hawaii and Australia.

In 1967–1968, two groups surveyed the southern sky. A group from Lawrence Livermore Laboratory (LLL) detected Cen X-3 twice and derived a rough location (Chodil *et al.*, 1967). Figure 1.5 shows data from one of these flights. However, a group from Leicester observed twice and did not see it (Cooke & Pounds, 1971).

In the late 1960s it was no easy task to build detectors, calibrate them, ensure that they survived the quick but hazardous trip into space, and know where they were pointed. People took pride in their ability to distinguish real sources from the background and expected the source population to be more or less steady, like the stars. In the case of Cen X-3, both groups secretly suspected that the other had not interpreted the data properly. In truth, all these observations were carefully done and correctly interpreted. The source is highly variable. To a small detector, sometimes it appears above background and sometimes not. Furthermore, such variability is a common characteristic of most bright X-ray sources. It took a while for people to believe this.

The Uhuru observations of Cen X-3, made in 1971, were spectacular (Giacconi *et al.*, 1971; Schreier *et al.*, 1972). The X-ray observations alone determined the nature of the source.

The first surprise was the observation of a regular periodicity of 4.84 seconds in X-ray flux from Cen X-3. The modulation was high and the pulsations were easily seen during a single scan across the source. Only a rotating neutron star could produce such rapid pulsations. The period was measured accurately and it was soon discovered that the period varied slightly with time. After several days of data were collected, these variations were recognised as a Doppler shift. The neutron star was moving in a circular orbit with a period of only 2.09 days. As icing on the cake, the X-rays were observed to disappear completely for 11 hours at

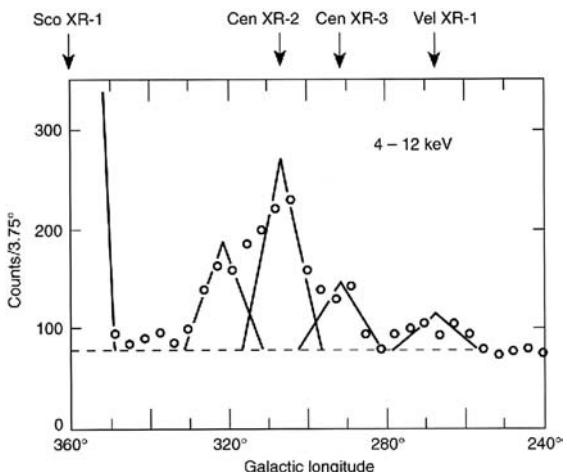


Fig 1.5 Counting rate from a rocket-borne proportional counter flown in May 1967. A slit collimator restricted the field of view to be $10^\circ \times 30^\circ$. The spinning payload caused the detector to scan a band of the sky repeatedly. By comparing the observed count rate with the expected triangular response to a point source, sources could be located. Sco X-1 was very bright and is off scale at 360° . The next strongest source at this time was Cen X-2, a transient. The source Cen X-3 was first seen in the data shown here. Note the difficulty of determining source positions in crowded regions (from Chodil *et al.*, 1967).

regular 2.09 day intervals. The source was in an eclipsing binary system.

Here then was a rapidly spinning neutron star, probably emitting X-rays from the near-vicinity of one of its magnetic poles. It orbits a bright B0Ib star, Krzeminski's star (named after the person who identified the optical counterpart). Energy to power the X-ray source comes through accretion of material supplied by the supergiant companion. This matter is captured by the strong gravitational field of the neutron star. It acquires enough energy in the fall to the surface to both heat material to the high temperature required for X-ray emission and to supply the observed luminosity (see Chapter 11).

The other bright X-ray sources in the plane of our Galaxy were first detected in early rocket surveys (e.g. Figures 1.5 and 1.6). Most were found by groups at Lockheed (Fisher *et al.*, 1968), MIT (Bradt *et al.*, 1968) and NRL (Friedman *et al.*, 1967). The sources are mostly accretion-powered binaries, in which a normal star and a compact star are locked in a close orbit. Some, like Cen X-3,

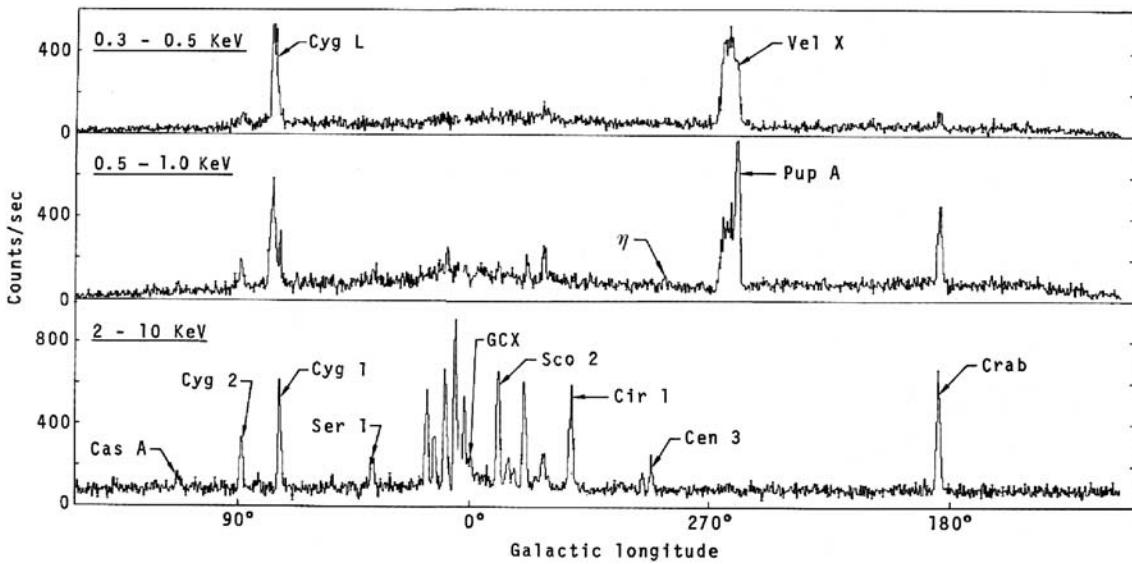


Fig 1.6 The entire Milky Way as surveyed with rocket-borne proportional counters in May 1970, May 1971, and October 1972 (Seward et al., 1972). Collimation was $1.3^\circ \times 20^\circ$. Data from the three flights have been combined to show counting rate as a function of galactic longitude in three energy bands. There are no soft X-rays observed from the cluster of bright sources around the galactic centre. Intervening gas absorbs the soft X-rays. The nearby supernova remnant Vela XYZ is clearly soft and extended. These data were taken using the payload shown in Colour Plate I, which was recovered and refurbished after each flight (figure available from FDS).

consist of a neutron star and a bright O star. The optical identifications of these were quickly made. Because the O stars are physically large, eclipses of the X-ray source associated with the orbiting neutron star are not unusual. Other sources consist of dim late-type stars orbiting close to a neutron star. These optical counterparts are faint and difficult to identify. The accretion-powered sources are the most luminous in our galaxy. Some have X-ray luminosities, $L_x \approx 10^{38}$ erg s $^{-1}$.

Some bright sources were found by Uhuru to be within globular clusters. Clark and colleagues (1975) found more with the third Small Astronomy Satellite (SAS 3) and pointed out that this was an unusual situation. The sources occur with much higher frequency than predicted by calculations based on the ratio of stars to X-ray sources in our galaxy (Clark et al., 1975). The high stellar density in globular clusters is clearly favourable for the formation of these exotic binary systems (see Chapters 11 and 12).

In 1973, soft X-rays from SS Cygni were discovered by Rappaport et al. (1974). SS Cyg is a

cataclysmic variable (CV) and is one of the brightest and nearest of this class. It has irregular outbursts during which the star brightens from its normal 12th to 8th magnitude. SS Cyg has been monitored by the American Association of Variable Star Observers (AAVSO) since 1896. It has an outburst about every 2 months and is called a *dwarf nova*. Many CVs are now known to be X-ray sources. They are accreting binary systems consisting of a low-mass normal star and a white dwarf (see Chapter 10).

Supernova remnants are bright X-ray sources, such as the first to be detected, the Crab Nebula. The X-ray luminosity of most remnants is 10–100 times less than that of the Crab Nebula and the spectrum is soft, so absorption in interstellar gas is more severe. Nevertheless, the closer remnants were easily detected and positively identified by their spatial extent (e.g. Cygnus Loop and Vela XYZ as seen in Figure 1.6) (Grader et al., 1970) or by the spatial coincidence with a non-thermal radio source such as Cas A (Gorenstein et al., 1970) (see Chapter 8).

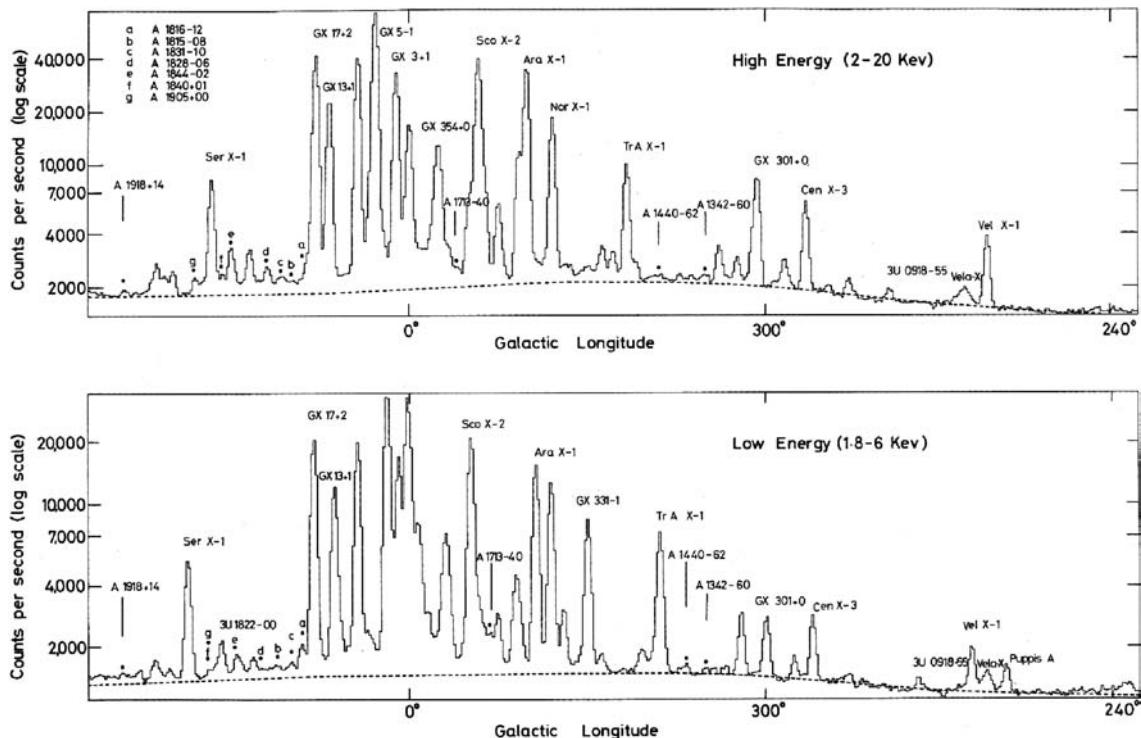


Fig 1.7 An Ariel 5 scan of the central half of the galactic plane. Two detectors scanned the sky, each with $0.7^\circ \times 10^\circ$ field of view. The two collimators were inclined at different angles to aid in source location, so given sources do not appear at identical longitudes in this figure. Note the improvement in the ability to detect weak sources. (Courtesy of K. Pounds.)

Another class of sources are stars, binary perhaps, but without compact companions. The first indication of strong coronal emission from stars was obtained in April 1974 (Catura *et al.*, 1975). The X-ray luminosity was 10 000 times the X-ray luminosity of the Sun. The detection occurred by accident when the rocket-borne instruments were pointed at Capella to calibrate star sensors included in the payload for an accurate measure of pointing direction. Shortly afterwards, in October 1974, X-ray emission from a second star, the flare star YZ Canis Minoris, was observed with the ANS by Heise *et al.* (1975).

Because a bright star in an error box was a very tempting identification, false claims of X-ray detection of bright stars were not uncommon. In spite of this, it was soon evident that the coronal X-ray emission of many active stars was considerably more intense than that of our Sun (see Chapter 6).

As the sensitivity of observations increased, other sources were discovered that were not in our Galaxy. The first extragalactic source discovered was the active galaxy M87. The observation was made by Byram *et al.* (1966) with a rocket launched April 1965. In 1971, Uhuru added many quasars, active galaxies, and clusters of galaxies (Giacconi, 1974). Thus, the individual X-ray source populations were recognised as sources and identified.

Large fractions of the sky were surveyed by the first satellites devoted to X-ray astronomy. After Uhuru, SAS 3 and Ariel 5, a few hundred sources had been catalogued (e.g. Figure 1.7). The first satellite/observatory specifically designed for an all-sky survey was HEAO-1 in 1979, which used an array of large-area proportional counters. The result, shown in Colour Plate 2, was a catalogue with limiting sensitivity of $0.003 \text{ photons cm}^{-2} \text{ s}^{-1}$ containing 842 sources (Wood *et al.*, 1984). Ten

years later, ROSAT mapped the sky for the first time using an imaging telescope and low-background detector. The threshold of this, the most sensitive X-ray all-sky survey to date, was 1.5×10^{-4} photons $\text{cm}^{-2} \text{ s}^{-1}$, and the first version of the ROSAT catalogue contained more than 18 000 sources, both galactic and extragalactic (Voges *et al.*, 1999). Colour Plate 3 shows the ROSAT all-sky survey.

Flux and luminosity

Fluxes quoted are measured at the top of the Earth's atmosphere. To give an intuitive feeling for the X-ray brightness of a source, fluxes in this chapter have been quoted in units of photons $\text{cm}^{-2} \text{ s}^{-1}$. The counting rate of an X-ray detector, $C = F_p \epsilon A$, where F_p is photon flux, ϵ is detector efficiency and A is detector area integrated over the energy range of the detector. Because detector efficiencies usually ranged from 0.1 to 1.0, and detector areas from 100 to 1000 cm^2 , the counting rate of early X-ray detectors was ~ 100 times the photon flux quoted.

To be more precise, we should specify an energy flux ($\text{ergs cm}^{-2} \text{ s}^{-1}$) and the exact energy range covered. The observed flux is a measure of the brightness of a source. The intrinsic luminosity, L , is related to the flux, F , through the square of the distance to the source, d . Thus, $L = 4\pi d^2 F$. As a matter of interest, in the range 0.2–10 keV, one of the most luminous X-ray sources known is the quasar PKS 2126-150, at a red shift of 3.27 and with $L_x = 5 \times 10^{47}$ erg s^{-1} . One of the least luminous extra-terrestrial X-ray sources detected is the Earth's Moon with $L_x = 7 \times 10^{11}$ erg s^{-1} , a range of physical processes that produce X-ray emission varying by 36 orders of magnitude.

In a more selective mode of operation, X-ray telescopes have accomplished deep surveys of small regions of the sky. The Einstein and ROSAT deep-survey detection thresholds were $\approx 3 \times 10^{-5}$ and $\approx 1 \times 10^{-5}$ photons $\text{cm}^{-2} \text{ s}^{-1}$, respectively. The XMM and Chandra limits are 100 times fainter again or ≈ 3 photons $\text{m}^{-2} \text{ hr}^{-1}$. This is 1 billion times fainter than Sco X-1, the brightest source in

the sky. (Actually, at energies above ~ 1 keV, transient sources up to twice as bright as Sco X-1 have been observed. Below ~ 1 keV, as you know, the Sun is the brightest source in the sky.)

References

Observation

- Bradt, H., S. Naranan, S. Rappaport & G. Spada (1968) Celestial positions of X-ray sources in Sagittarius, *ApJ* **152**, 1005.
- Bowyer, S., *et al.* (1964) X-ray sources in the galaxy, *Nature* **201**, 1307.
- Byram, E., T. Chubb & H. Friedman (1966) Cosmic X-ray sources, galactic & extragalactic, *Science* **153**, 1527.
- Catura, R. C., L. W. Acton & H. M. Johnson (1975) X-ray emission from Capella, *ApJ* **196**, L47.
- Chodil, G., *et al.* (1967) Spectral and location measurements of several cosmic X-ray sources including a variable source in Centaurus, *Phys. Rev. Letters* **19**, 681.
- Clark, G. W. (1965) Balloon observation of the X-ray spectrum of the Crab Nebula above 15 keV, *Phys. Rev. Letters* **14**, 91.
- Clark, G. W., T. H. Markert & F. K. Li (1975) Observation of variable sources in globular clusters, *ApJ* **199**, L93.
- Cooke, B. A. & K. A. Pounds (1971) Further high sensitivity X-ray sky survey from the Southern Hemisphere, *Nature Phys. Sci.* **229**, 144.
- Fisher P. C., *et al.* (1968) Observations of galactic X-ray sources, *ApJ* **151**, 1.
- Friedman, H., E. T. Byram & T. A. Chubb (1967) Distribution and variability of cosmic X-ray sources, *Science* **156**, 374.
- Giacconi, R., H. Gursky, F. Paolini & B. Rossi (1962) Evidence for X-rays from sources outside the Solar System, *Phys. Rev. Letters* **9**, 439.
- Giacconi, R., *et al.* (1971) Discovery of periodic X-ray pulsations in Centaurus X-3, *ApJ* **167**, L67.
- Giacconi, R., *et al.* (1974) The Third UHURU Catalog of X-Ray Sources, *ApJ Supp* **27** 37.
- Gorenstein, P., E. Kellogg & H. Gursky (1970) X-ray characteristics of three supernova remnants, *ApJ* **160**, 199.
- Grader, R. J., R. W. Hill & J. P. Stoering (1970) Soft X-rays from the Cygnus Loop, *ApJ* **161**, 145.
- Gursky, H., R. Giacconi, F. Paolini & B. Rossi (1963) Further evidence for the existence of galactic X-rays, *Phys. Rev. Letters* **11**, 530.

- Gursky, H., et al. (1966) A measurement of the location of the X-ray source Sco X-1, *ApJ* **146**, 310.
- Heise, J., et al. (1975) Evidence for X-ray emission from flare stars observed by ANS, *ApJ* **202**, L73.
- Hill, R. W., R. J. Grader & F. D. Seward (1968) The soft X-ray spectrum of Sco XR-1, *ApJ* **154**, 655.
- Rappaport, S., et al. (1974) Possible detection of very soft X-rays from SS Cygni, *ApJ* **187**, L5.
- Schreier, E., et al. (1972) Evidence for the binary nature of Centaurus X-3 from Uhuru X-ray observations, *ApJ* **172**, L79.
- Seward, F. D., et al. (1972) Distances and absolute luminosities of galactic X-ray sources, *ApJ* **178**, 131.

Catalogues, reviews

- Hirsch, R. F. (1983) *Glimpsing an Invisible Universe*, Cambridge: Cambridge University Press.
- Tucker, W. H. & R. Giacconi (1985) *The X-ray Universe*, Cambridge, MA: Harvard University Press.
- Voges, W., et al. (1999) The ROSAT all-sky survey bright source catalogue, *A&A* **349**, 389.
- Wood, K. S., et al. (1984) The HEAO A-1 X-ray source catalog, *ApJ Supp* **56**, 507.

Chapter 2

X-ray emission and interaction with matter

2.1 Astrophysical mechanisms for generating X-rays

There are three radiation processes – thermal, synchrotron and blackbody – that are the dominant mechanisms for producing X-rays in an astronomical setting, and whenever high-energy electrons are present, we must add inverse Compton scattering of microwave background photons into the X-ray regime. The spectral signature of each process is unique and is therefore one of the first clues to the nature of an unknown X-ray source. If the spectrum can be measured with high resolution over a broad energy band, then usually both the emission process and the physical conditions within the source can be deduced.

2.1.1 Thermal emission from a hot gas

Consider a hot gas of low enough density that it can be described as thin and transparent to its own radiation. This is not difficult to achieve for X-rays. At temperatures above 10^5 K, atoms are ionised, and a gas consists of positive ions and negative electrons. Thermal energy is shared among these particles and is transferred rapidly from one particle to another through collisions. Indeed *thermal equilibrium* means that the average energy of all particles is the same and is determined only by the temperature. When an electron passes close to a positive ion, the strong electric force causes its trajectory to change. The acceleration of the electron in such a collision causes it to radiate electromagnetic energy, and

this radiation is called *bremsstrahlung* (literally, ‘braking radiation’).

Electrons in thermal equilibrium have a well-determined distribution of velocities (called *Maxwellian* after the physicist James Clerk Maxwell), and the radiation from such electron-ion collisions is a continuum with a characteristic shape determined only by the temperature. This is *thermal bremsstrahlung*. The higher the temperature, the faster the motion of the electrons and hence the higher the energy of the photons in the bremsstrahlung radiation. For temperatures above 1 million degrees, these photons are predominantly X-rays.

The thermal bremsstrahlung spectrum falls off exponentially at high energies and is characterised by the temperature T . The intensity, I , of the radiation at energy E is given by

$$I(E, T) = AG(E, T)Z^2 n_e n_i (kT)^{1/2} e^{-E/kT},$$

where k is Boltzmann’s constant and G is the *Gaunt* factor, a slowly varying function with value increasing as E decreases. The form of this spectrum can be seen in Fig. 2.1. Note that the intensity is proportional to the square of the charge of the positive ions, Z , and the product of the electron density, n_e , and the positive ion density, n_i . A is a constant.

In a hot gas, X-ray line emission is also an important source of radiation. The elements heavier than hydrogen are not completely ionised, except at very high temperatures ($\geq 5 \times 10^7$ K). When a fast electron strikes an ion with bound electrons, it often transfers energy to that ion, causing a transition to a higher energy level. The

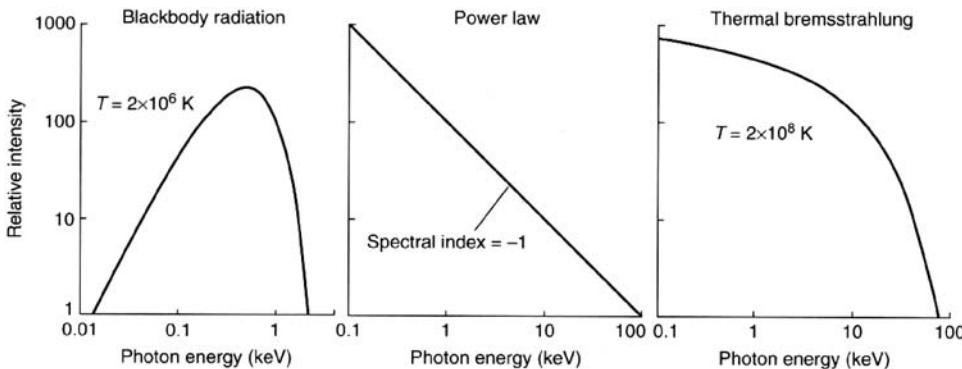


Fig 2.1 Spectral forms expected from different astrophysical processes.

ion is left in an excited state which lasts only briefly. The ion decays to its ground state by radiating photons of energy characteristic of the spacing of energy levels through which the excited electron passes. This radiation appears as spectral lines with energies determined by the radiating ion species.

Radiation from a thermal gas is thus a blend of thermal bremsstrahlung and line radiation (other processes also make small contributions, but X-ray diagnostics usually rely on these two). For a gas of ‘cosmic’ composition (which, for every 10 000 atoms of hydrogen, contains 850 atoms of helium and 16 atoms of carbon, oxygen and heavier elements), at temperatures below 10^6 K, most of the energy is radiated as UV lines. At 2×10^6 K, half the energy is radiated as soft X-rays; at 10^7 K, all the energy is radiated as X-rays, half in lines and half as thermal bremsstrahlung. At 5×10^7 K, almost all the ions have been stripped of their bound electrons, and almost all the energy is radiated in the X-ray continuum. Figure 2.2 shows how ion state depends on temperature.

Thus, by measuring X-ray spectra, the shape of the continuum and/or the presence of lines can identify the origin as a hot gas. The temperature of the gas can be calculated from the particular lines present and from the shape of the high-energy end of the bremsstrahlung continuum. The strength and energies of the lines reveal the elemental composition of the gas.

The power, P , radiated by 1 cm^3 of cosmic-composition gas at uniform temperature is given by the cooling curve shown in Figs. 2.3 and 2.4.

The luminosity of a volume, V , of this gas is $L = P(T)n_e n_i V$. If the distance of a diffuse X-ray source is known, the volume of X-ray-emitting gas can usually be estimated and the gas density derived from this expression.

The quantity $n_e n_i V$ is called the *emission measure*. Astrophysical sources, such as stellar coronae, are seldom isothermal, and a source with volumes of gas at different temperatures is described by a *differential emission measure* (DEM). One definition is $\text{DEM}(T) = n_e n_i \Delta V / \Delta T$. Although the definition varies a bit from paper to paper, the DEM is always used to give the temperature distribution of the emitting plasma. An example is shown in Chapter 6.

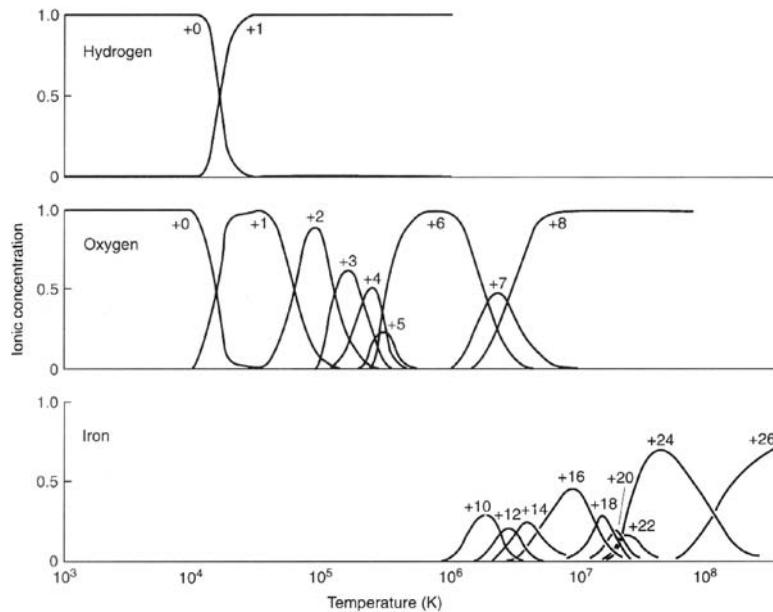
2.1.2 Synchrotron radiation from relativistic electrons

A fast electron traversing a region containing a magnetic field will change direction because the field exerts a force perpendicular to the direction of motion. Because the velocity vector changes, the electron is accelerated and consequently emits electromagnetic energy.

This is called *magnetic bremsstrahlung* or *synchrotron radiation* (after radiation observed from particle accelerators by that name). The frequency of the radiation depends on the electron energy, the magnetic field strength B and the direction of motion relative to the field.

In an astrophysical setting, the magnetic field can be somewhat aligned, but particle velocities are expected to be isotropic, so the observed

Fig 2.2 Ionic equilibrium of hydrogen, oxygen and iron in a thin gas as a function of temperature. At a temperature of 3000 K, all H atoms have a bound electron, whereas by 16 000 K, half the atoms are ionised. Above 30 000 K, no electrons are bound to H nuclei. As temperature rises, more and more electrons are stripped from the heavier atoms. These curves are recalculated whenever new atomic physics data are available, and recent calculations predict slightly different temperature dependences for Fe than shown here (taken from calculations of D. Cox and C. Jordan; see Jordan, 1969).



spectrum depends only on B and the energy spectrum of the electrons. The usual spectral form assumed for the electrons is a power law, and if this is so, then the spectrum of the resulting synchrotron radiation is a power law also. Indeed, when an observed spectrum is a power law over a reasonably large energy range, it is usually taken as a strong indication that the source is emitting synchrotron radiation. If the magnetic field is aligned, the radiation will be polarised, and observed polarisation is usually proof of synchrotron emission.

The form of the power-law spectrum is simply $I(E) = AE^\gamma$, where A is a constant and γ is the spectral index. The more negative the value of γ , the softer is the spectrum. The spectrum shown in Fig. 2.1 has $\gamma = -1$.

The power radiated as synchrotron radiation is proportional to B^2E^2 and the average photon energy to BE^2 . In an astronomical setting, synchrotron X-rays indicate the existence of very energetic electrons. Two examples of such synchrotron radiation are (1) radio emission from shell-like supernova remnants where the magnetic field strength is $\approx 7 \times 10^{-5}$ Gauss and the electrons which radiate radio waves have energies of about 1 GeV; (2) the central region of the Crab Nebula where the magnetic field is ≈ 10 times

stronger than this. Radiation is emitted over most of the electromagnetic spectrum, and the electrons which produce synchrotron X-rays here have energies of about 10^4 GeV, a few ergs each! Furthermore, because these electrons lose energy rapidly, there must be a continuous injection of fresh electrons, which is achieved by their being accelerated in the rapidly spinning pulsar magnetosphere (see Chapter 9).

The derivation of synchrotron emission assumes that electrons travel in a uniform magnetic field and at an angle to the field lines. There is no radiation when particles move parallel to the field lines. The field lines around a neutron star, for example, are not straight, and an electron moving along one of these lines will be accelerated and will radiate. This special case of synchrotron radiation is called *curvature radiation*.

2.1.3 Inverse Compton X-rays from relativistic electrons

An ultra-relativistic electron can collide with photons and produce X-rays. In the rest frame of the electron, there is Compton scattering, and the photon transfers energy to the electron. In the laboratory frame, the scattered photons are Doppler shifted to high energies, and the process is called *inverse Compton scattering*.

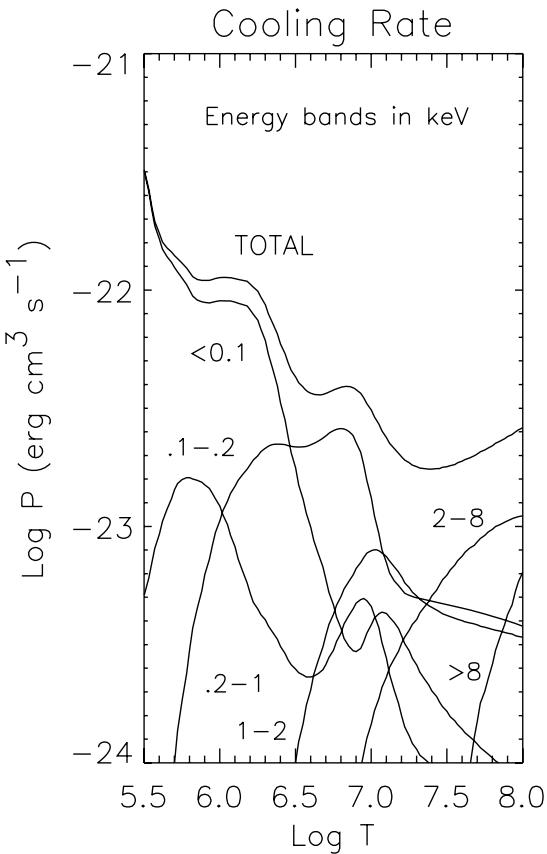


Fig 2.3 Power radiated from a low-density plasma as a function of temperature in several X-ray wavebands. Each curve is labelled with the energy range in keV. The material is in thermal equilibrium, and abundances are cosmic. The counter-intuitive result that the emission is stronger at lower temperatures is because of the contribution of emission lines, which disappear at higher temperatures as the atoms become completely ionised (figure from John Raymond; update of Raymond et al., 1976).

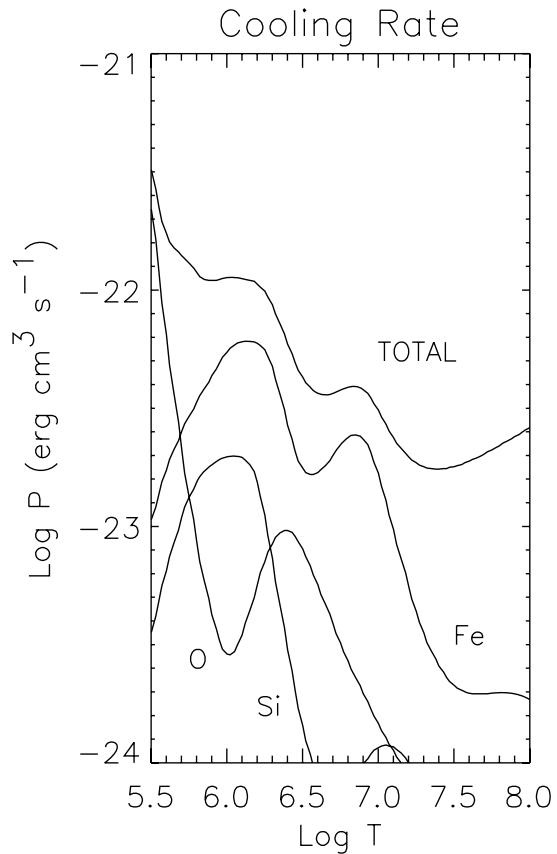


Fig 2.4 Power radiated from a low-density plasma as a function of temperature showing contributions because of collisional excitation from three elements (figure from John Raymond; update of Raymond et al., 1976).

The total energy of a relativistic particle is γmc^2 , where γ is the Lorentz factor, $[1 - (v/c)^2]^{-1/2}$. If the photon environment has energy density u_{rad} and photon energy is $h\nu$, the inverse Compton power radiated is proportional to $\gamma^2 u_{\text{rad}}$ and the scattered photon has energy $\gamma^2 h\nu$. Thus cosmic-ray electrons with $\gamma = 1000$ will make keV X-rays when colliding with 10^{-3} eV cosmic microwave background (CMB) photons and MeV γ rays colliding with starlight. Because the CMB pervades all space, inverse Compton scattering is a serious energy loss for any ultra-relativistic

particles traveling intergalactic distances. The process is also important in quasar jets where u_{rad} is high and high-energy particles are abundant.

2.1.4 Blackbody radiation from starlike objects

A ‘black’ surface completely absorbs any radiation incident upon it. Reflectivity is zero. It is furthermore a law of nature that the surface must not only absorb but also emit radiation. The spectrum radiated is a well-defined continuum with peak emission at an energy dependent only on the temperature, T . The higher the temperature, the more energetic the photons. A familiar example is the electric heating element on a stove or hot plate. As it is heated

it first glows a deep red and then, as the temperature increases, becomes orange and almost yellow.

The form of the spectrum is given by

$$I(E, T) = 2E^3[h^2c^2(e^{E/kT} - 1)]^{-1},$$

where h is Planck's constant and c is the speed of light.

The stars radiate as blackbodies with temperatures from 2500 K (red dwarf) to 40 000 K (O star). Although strongly modified by the stellar atmospheres the spectra retain the overall gross shape imposed by the blackbody emission process. A newly formed neutron star is expected to have a very hot surface. If the surface temperature is 10^6 K or higher, it will emit blackbody radiation with photons in the X-ray range. The spectra of isolated neutron stars do have this form, but usually, strong magnetic fields at the surface cause complications.

2.1.5 Bombardment

This last process applies to much of Chapter 4. In the laboratory, X-rays are produced by bombarding a target with a beam of electrons. In the solar system, planets, moons and comets are bombarded with particles from the Sun and from planetary magnetospheres. The target emits bremsstrahlung X-rays and fluorescent X-rays with energies characteristic of the target and of any bombarding ions. Solar X-rays also bombard cold objects which emit scattered and fluorescent X-rays.

2.2 Interaction of X-rays with matter

When a beam of X-rays passes through a slab of material, some photons disappear from the beam. There is no gradual loss of energy, as for a beam of charged particles. X-rays collide with individual atoms, and the photon energy is deposited in surrounding material. The transmitted beam has fewer photons, but the energy of each transmitted photon has not changed. For a slab of thickness x , the intensity of the transmitted beam is reduced by a factor $e^{-\mu x}$, and μ is called the *absorption*

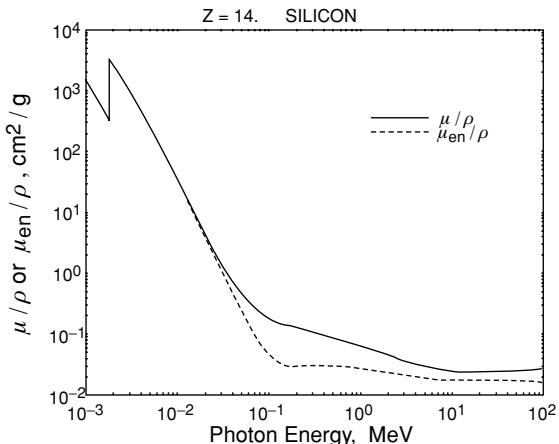


Fig 2.5 X-ray attenuation coefficient for Si as a function of energy from 1 keV to 100 MeV. The solid curve shows probability of X-ray interaction. The dashed line shows the fraction of the photon energy that is transferred to the kinetic energy of charged particles. Photoelectric absorption dominates at low energies and decreases rapidly as energy increases. At 1.839 keV, ejection of a K-shell electron is possible. At ≈ 30 keV, Compton scattering becomes important. Above ≈ 20 MeV, pair production is important (from Hubbell & Seltzer, 1996).

coefficient. Working with solid materials on Earth, it is customary to use the density ρ to define a mass-absorption coefficient μ/ρ and to measure thickness in grams per centimeter squared. In astronomy, thickness is measured as the number of atoms N_H in a 1-cm² column between the observer and source, and the absorption coefficient becomes a cross section, usually written as σ . Hence the absorbing factor is usually written as $e^{-\sigma N_H}$.

X-rays interact with matter in several ways. In the energy range 0.1–10 keV, the main interaction is the photoelectric effect. An atom absorbs the photon and ejects an electron with energy equal to that of the photon minus the binding energy of the electron, i.e. $E_{el} = E_{ph} - E_b$. The photoelectric cross section varies with energy as $\sim Z^3 E^{-3}$, so absorption is greatest at low energies and in high-Z materials. For a given material, as photon energy increases, μ decreases, but when the energy becomes high enough to free one of the more tightly bound electrons, there is an abrupt increase in μ . Figure 2.5 shows the absorption coefficient energy dependence of one element

(arbitrarily chosen to be Si). In the energy range 0.1–10 keV, X-ray interstellar absorption curves show ‘edges’ corresponding to the binding energies of electrons in the K and L shells of light elements. Figures 5.1 and 5.2 show these expected effects in the ISM.

X-ray photons also scatter from individual electrons within the atom. This process is called *Compton scattering* and is not important at low energies. In this process, the energy of the incident X-ray is shared between the scattered photon and the electron. The electron energy depends on the scattering angle and is maximum when the photon is scattered back in the incident direction. The maximum energy transferred to the electron is appreciable at high energies and very small at low energies. A 100-keV X-ray can transfer up to 28 per cent of its energy to a Compton electron. A 10-keV photon can transfer up to 4 per cent, and a 1-keV X-ray can transfer up to only 0.4 per cent (4 eV). (A third process becomes possible above 1.02 MeV. Here all the photon energy can go into production of an electron-positron pair. This is important for γ -ray detection but far above the energies considered in this book.)

So X-ray detectors in the range 0.1–10 keV respond to energy deposited by photoelectrons in the active volume of the instrument. At higher energies Compton electrons can also play a role. The next chapter concerns details of such detectors. X-rays also scatter from atoms as a whole with no energy loss. This occurs in diffraction gratings where the wave nature of X-rays is utilised for high-resolution energy measurements and at low incidence angles when X-rays are scattered from smooth surfaces, as in X-ray telescope mirrors, as discussed in the next chapter.

References

- Hubbell, J. H. & S. M. Seltzer (1996) NIST tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients, <http://physics.nist.gov/PhysRefData/XrayMassCoef/ElemTab/z14.html>.
- Jordan, C. (1969) The ionization equilibrium of elements between carbon and nickel, *MNRAS* **142**, 501.
- Raymond, J. C., D. P. Cox & B. W. Smith (1976) Radiative cooling of a low-density plasma, *ApJ* **204**, 290.

Chapter 3

Tools and techniques

3.1 X-ray detectors

The first instruments used for X-ray astronomy were developed originally for the detection of charged particles and γ rays emitted by radioactive material. These detectors respond to energy deposited by photoelectrons and, for higher energies, Compton electrons (discussed in Chapter 2). A fast electron creates a track of ionised material in the active volume of the detector. The detector collects either this charge or light from recombination of the ions. Electronic circuits then amplify this signal and record the time and amplitude of the event.

3.1.1 The proportional counter

The proportional counter is not only an efficient X-ray detector but also measures the energy of every photon detected. It was the workhorse of early cosmic X-ray observations and is still being used in modern instruments. However, the modifications necessary to adapt the simple laboratory counter to an X-ray detector capable of operating in the harsh environment of space were challenging.

The detector must have a large area to collect photons from weak cosmic sources and obviously a window thin enough to transmit X-rays. Yet the window has to be strong enough to keep the gas inside the detector from leaking into the near-vacuum of space and well supported to withstand the force of the gas pressure inside the detector. Many an early observation was lost by the failure of detector windows during rocket ascent out of

the atmosphere and upon first exposure to space. A typical rocket-borne proportion counter, such as that shown in Fig. 3.1, had a window of 0.1 mm Be or of even thinner plastic, such as Formvar, and an area of 100–300 cm². The detectors were filled with a noble gas-methane mixture (A, Ne and Xe were used) at slightly more than 1 atmosphere pressure.

When an X-ray photon enters the detector through the thin window, it is absorbed by atoms of gas contained inside the counter. The resultant photoelectrons produce short tracks of ions and electrons inside the gas. These electrons are attracted towards and drift to the vicinity of the central wire or anode (maintained at about +2000 volts), where they cause further ionisations by colliding with other atoms. An avalanche of

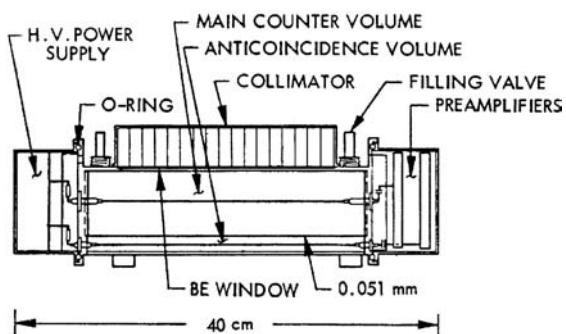


Fig 3.1 A proportional counter used in early sounding rocket observations. A slit collimator restricts the field of view, and the central volume is surrounded on three sides (one shown here) by guard counters to reject events caused by cosmic rays (from Chodil *et al.*, 1967).

Box 3.1 | Rejection of charged-particle events

Above the atmosphere a detector is bathed in a flux of charged particles: high-energy cosmic rays and sometimes trapped or precipitating particles. These produce background events in the detector which can overwhelm the weak signals from astronomical sources unless precautions are taken.

An X-ray photon can only enter the detecting volume by passing through the collimator and the thin entrance window. Cosmic rays, however, can pass completely through the walls of an X-ray detector and even through the spacecraft carrying it! In doing so, they deposit a line of charge through the counter gas which is then detected in the same way as that produced by real X-rays. A second detector, called a *guard counter*, is usually placed close to the X-ray detector. Most cosmic rays traversing the X-ray detector will also pass through the guard counter. Any event detected simultaneously in both volumes can be discarded by on-board electronic logic. The guard counter is said to be in *anti-coincidence* with the main X-ray counter or is being used as a *veto* system. Guard counters can be built into the same chamber as the X-ray detector so that they use the same gas supply. Sometimes several guard counters surround the X-ray detector on all sides.

An X-ray photoelectron makes a track ~ 1 mm in length in the detector gas. This produces a short (≤ 10 ns) rise-time pulse at the anode. A cosmic ray produces a much longer trail (several centimeters), and the charge therefore arrives at the anode

spread over a much longer period of time (> 100 ns). These times can be measured electronically for every event, and only short-rise-time X-ray events are admitted by the electronics. This technique for charged-particle event rejection is called *pulse rise-time discrimination*. The height (or amplitude) of the pulse can also be used as a discriminant to reject charged particles as they usually deposit considerably more energy in the detector than do X-rays.

Thus more than 90 per cent of cosmic-ray-induced events can be rejected. There are also low energy protons and electrons moving in the Earth's magnetic field that can penetrate the window and deposit a few keV in the detector. Magnets built into the collimator can sweep these into the collimator sides and prevent them from entering the window.

The flux of particles in the radiation belts surrounding the Earth is high, and astronomical detectors risk being severely damaged or destroyed if they are activated within that region. The resultant high count rates swamp the high voltage anodes, causing serious breakdown problems. Even near-Earth orbits can pass through a region known as the *South Atlantic Anomaly*, in which the charged particle flux is high enough to damage detectors. Precautions must be taken to operate a proportional counter on a satellite. Even rocket observations made at lower altitude can be troubled by precipitating electrons.

electrons occurs close to the anode, and the net result is a burst of electrons that are collected on the central wire. The avalanche amplifies the original signal (the number of electron-ion pairs created by the photoelectron) in a linear way. Events from the detector output, sorted by size, form a *pulse-height spectrum* which is clearly related to the energy of the incident X-rays (through the photoelectrons). Events from UV photons are discriminated against by rejecting all pulses below a certain threshold.

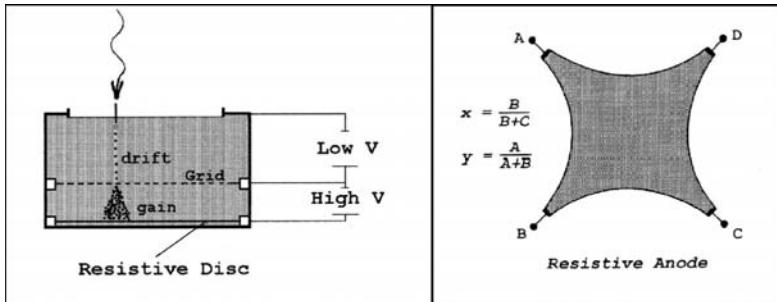
In a Geiger counter (as used in the original AS&E experiment) the voltage is higher, and the avalanche becomes a discharge which gives no

more information than the fact that the event occurred. However, in the proportional counter the central wire is maintained at a slightly lower voltage so that the number of electrons collected by the anode is proportional to the energy of the incident X-ray photon.

3.1.1.1 Imaging proportional counters

Several techniques can be used to record the position of an event within a proportional counter. If the anode is a single wire, but made of resistive material, the position of the event along the length of the wire can be determined by the relative size of the pulse measured at the two ends of

Fig 3.2 (left) A diagram of the position-sensitive proportional counter used on EXOSAT. The avalanche occurs close to the cathode so the charge is confined to a particular spot on the plate. (right) Signals from the corners of the plate are combined to measure the location of the event.



the wire. When used at the focus of a telescope, the counter becomes two-dimensional. Positive ions from the localised avalanche travel to the cathode and deposit their charge in a limited area close to where the avalanche occurred.

By making the cathode a resistive plate (as illustrated in Fig. 3.2) and measuring the size of the pulse at the four corners, the location of the deposited charge can be determined. This system was used in one of the EXOSAT detectors. Another technique used in the Einstein imaging proportional counter (IPC) was a cathode made of crossed resistive wire grids.

3.1.2 Gas-scintillation proportional counter

The gas scintillation proportional counter (GSPC), used on several missions, is similar in operation to a conventional proportional counter. Instead of detecting the electron cloud produced when the X-ray photon enters the counter, the GSPC detects the optical flash or scintillation that occurs when the ionised atoms in the gas recombine (i.e. rejoin with an electron). This process is analogous to that of the crystal-based scintillation counter used for detecting energetic X-rays. The avalanche is dispensed with and the intrinsic energy resolution is better than a standard proportional counter; it is close to that of a solid-state detector but without the need for cooling to very low temperatures.

3.1.3 The scintillation counter

Proportional counters do not detect photons with energies above 20 keV efficiently. This is because the detecting volume is a gas and, at high energies, the gas becomes transparent. A thicker (i.e. more absorbing) detecting medium is needed, and high-Z material is an advantage.

The scintillation counter uses crystals of sodium iodide or caesium iodide, which can efficiently stop photons with energies up to several MeV. The photon energy is absorbed by an atom within the crystal, and some of this energy immediately reappears as a pulse of light or *scintillation*. A photomultiplier tube then detects the scintillation, thus registering the time of the event. The amount of light in the scintillation is proportional to the incident X-ray photon energy.

For rejection of events caused by cosmic rays passing through the detector, the principal crystal was often surrounded by a second scintillating material. This could be in a separate light-tight assembly with its own photomultiplier. Because it is easily machinable, plastic scintillator was often used for the cup-shaped guard detector. However, the light pulses from plastic scintillator have faster risetimes than those from NaI or CsI. The NaI X-ray detector can be surrounded with plastic scintillator and all viewed by a single photomultiplier. The electronics then measure the shape of each light pulse and rejects fast-risetime events which show that a charged particle has passed through the outside scintillator. This detector is called a *phoswich*, short for ‘phosphor sandwich’. Scintillation counters have been the workhorses of balloon-based high-energy X-ray astronomy. At altitudes of more than 40 km, observations have been made in the energy range 20–200 keV.

3.1.4 Channel electron multiplier and micro-channel plate

The channel electron multiplier is a small-diameter glass tube which has been treated to enhance secondary emission properties. The pore

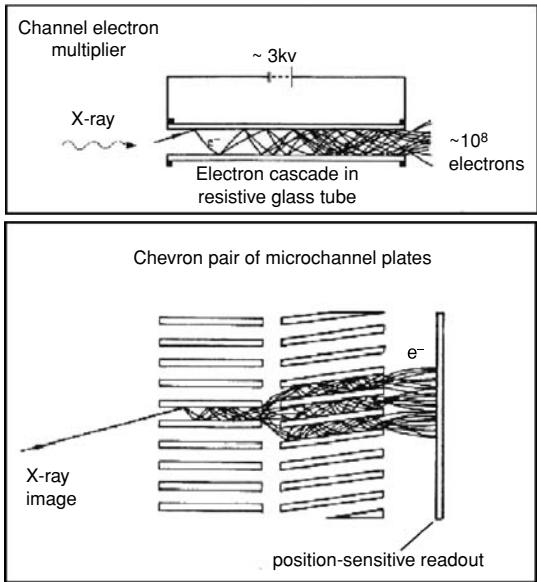


Fig 3.3 The micro-channel plate is a bundle of thousands of small glass tubes, each of which gives a large gain in the signal. A pair of plate assemblies prevents light from directly reaching the detector, and an image is transferred with high gain from the front surface to the detector.

diameter is $\approx 10\mu$. A high voltage is applied along the length of the tube. A photon striking the inside of the negative end ejects low-energy electrons which are accelerated down the tube. As they progress, they strike the wall and liberate more electrons. The original electrons become a cascade of as much as 10^8 electrons at the positive end. The negative end, or photocathode, is coated with CsI to improve the efficiency of X-ray photoelectron emission.

A micro-channel plate (or MCP) is a position-sensitive detector useful for recording images. It is simply a very large number of such tubes fused together by means of glass fibre technology. They have extensive military applications and are perhaps best known for their use in night-vision binoculars. A typical MCP of 25 mm diameter might have 3 million individual micro-channels, each $20\ \mu\text{m}$ across. It is capable of producing a gain of about 10^4 (i.e. a single electron entering a micro-channel will be amplified to yield 10^4 electrons at the other end, as shown in Fig. 3.3). The spacing of the channels is so small that the incident image is reproduced as an electron image

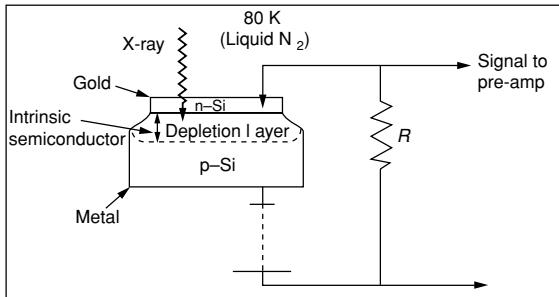


Fig 3.4 Diagram of the Einstein solid state spectrometer (SSS). The SSS was a reverse-biased junction that created a depletion layer as its detecting volume. An interacting X-ray produced ion pairs in this volume as in a proportional counter gas, but much more easily, and there was no avalanche. To reduce thermal noise, the whole system was cooled to 80 K , and a low noise amplifier was required.

which can be recorded by a variety of techniques. The Einstein and ROSAT high-resolution imager (HRI) detectors and the Chandra high-resolution camera (HRC) utilised micro-channel plates.

3.1.5 Solid state detector and charge-coupled devices

The solid state spectrometer (SSS) was first flown on Einstein and is shown schematically in Fig. 3.4. The incoming X-ray ionises atoms of material in the detector, but because the material is solid silicon and the applied voltage is low, there is no avalanche. The SSS collects the free electrons and measures the charge directly. This worked because the detector was solid state and was cooled to very low temperatures (80 K or -193°C) so that thermal noise was greatly reduced. Because each X-ray which deposited energy in the detector initially generated many more ion pairs than in the gas of a proportional counter, the resolution of the SSS was about three times better, yet with virtually no loss of efficiency.

Technology has advanced to the point where the SSS can be made much smaller and can be placed in arrays of $\sim 10^6$ detectors. These devices, the charge-coupled devices (CCDs), have been used in ground-based optical astronomy for more than 2 decades now, with revolutionary results. Their substantial gain in efficiency over previous devices has been directly responsible for many of the important discoveries of observational astronomy



Fig 3.5 The Chandra ACIS detector. The 2×2 ACIS-I and the 1×6 ACIS-S chip arrays are visible at the top (NASA/CXC 2010).

since ~ 1980 . The Chandra CCD is shown in Fig. 3.5. In this device, each chip of 1024×1024 pixels is 2.5 cm wide. The device is operated like a camera, and data are recorded during an interval of 3.2 s. If an X-ray is absorbed in a pixel, a charge proportional to the photon energy is deposited there. At the end of the recording interval, the device is read out one row at a time by circuitry

at one edge of the chip. After the first row is read out, charge in all pixels is shifted down one row and the second row is read out. This charge-transfer clocking process takes only 0.04 s to read all 1024 rows.

During normal operation, almost all pixels contain no charge, and only a few have charge deposited by a single X-ray. If the source is bright and the focus sharp, there can be two or more interactions in a single pixel during the 3.2 s recording interval. This charge is recorded as a single event having the summed energy of the interacting photons. This is *pileup*, a distortion of the spectrum which becomes worrying at count rates above ~ 0.01 counts $\text{pix}^{-1} \text{ s}^{-1}$. A very bright source will deposit many X-rays in a few central pixels during each recording cycle, and the charge can be so large that the event is rejected. An image of a very bright point source appears as a bright halo with a hole at the centre. Also, events are recorded during readout as the rows are clocked through the source location. This is the *charge transfer streak* or *CT streak*, which appears as a number of bright columns radiating from the source. Figures 3.6, 5.7 and 5.10 show bright CT streaks.

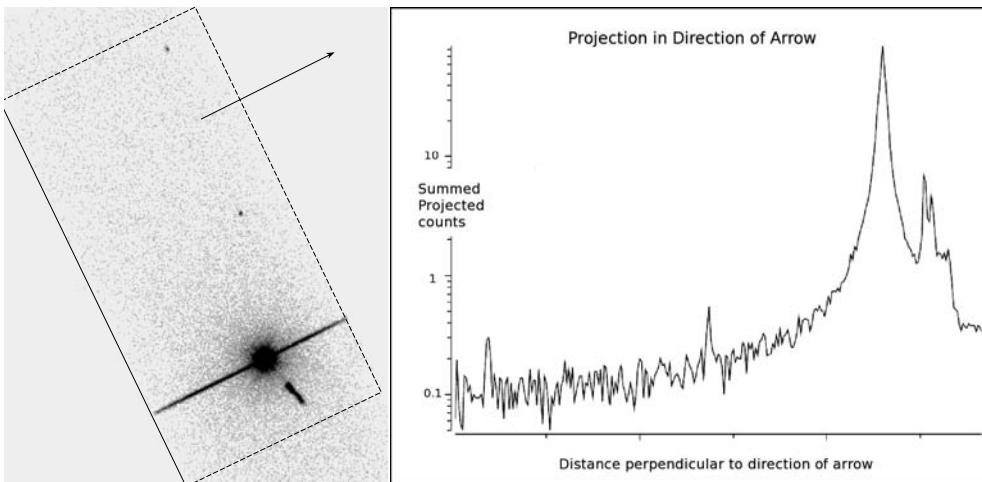


Fig 3.6 (left) A 40-ks Chandra ACIS image of the quasar 3C 273. The field covers $1.3' \times 2.6'$ of the sky. The quasar itself is a bright point source, overexposed in this presentation and with CT streak. There is a jet pointed southwest, separated from the quasar and with internal structure. The grey scale has been adjusted to show the broad, outer part of the mirror PSF. The width of the CT streak shows the size of the core of the PSF. (right) The image has been projected in the direction of the arrow, and the graph gives number of counts (log scale) in each column and shows structure within the jet. The core of the image is distorted by pileup and the CT streak. Note that the two faint sources are more easily visible in the image than in the projection.

3.1.6 Calorimeter

A remarkable and completely new device was developed and flown on the Suzaku mission in 2005. This was the quantum calorimeter, in which individual X-ray photons are absorbed by a crystal which is maintained at a temperature very close to absolute zero (≤ 0.1 K). The energy of the X-ray causes the temperature of the crystal to increase, and this is measured. The higher the energy of the photon, the greater the temperature increase; hence the device is a spectrometer. An energy resolution of 6 eV (~ 20 times better than the SSS!) was achieved (Kelley *et al.*, 2007).

Unfortunately, the Suzaku refrigerator failed, and no astronomy data were obtained with this detector. The method does show great promise, however, and calorimeters are planned for future missions. Much new technology is required. The temperature rise from an event is minute, and a very sensitive thermometer is needed. Some designs use a superconducting layer in thermal contact with the absorber to measure the temperature rise. Maintaining a low temperature for a long time is always a challenge. Furthermore, in a large array, a low temperature cannot be maintained if a wire connects each pixel to the warm environment. The rate of heat transfer must be minimised by using a pixel readout multiplexer built into the cold detector. As Table 3.1 shows, the gain in resolving power over the CCD is considerable,

and an extensive development effort will be well worthwhile.

3.1.7 Detection of polarisation

Polarisation measurements would be extremely valuable in the study of almost all types of cosmic sources. For example, scattering produces polarisation, and so an observation of an accretion disc corona might reveal how the material is distributed (see Chapter 11). Because emission in a magnetic field will produce polarised X-rays, we can learn details about the emission from pulsar-wind nebulae and from neutron star surfaces (see Chapter 9).

However, determining the level of X-ray polarisation is technically challenging, and there has been little progress through the history of X-ray astronomy. The first detectors utilised Thompson or Bragg scattering, in which the direction of the electric vector is preserved. Because this direction is always perpendicular to the direction of propagation, scattering of a linearly polarised beam is not isotropic. In particular, if the scattering angle is 90° , there is no scattering in the direction of the incident electric vector. This anisotropic scattering was the basis for the first detectors used to search for polarised X-rays from cosmic sources. These detectors comprised a scatterer and counters to register X-rays scattered at angles of $\approx 90^\circ$. When these detectors were rotated around the incident beam direction, the dependence of count rate on rotation angle was a measure of beam polarisation.

Because absorption is minimum in low-Z elements, the first rocket payload utilised Thompson scattering from Li blocks. A more sensitive crystalline-graphite Bragg instrument on the satellite OSO-8 actually managed to measure the polarisation of the Crab Nebula and set limits of a few per cent on polarisation from several bright accretion-powered binaries. To date, the only polarisation measure of X-rays from an extra-solar source is that of the Crab Nebula, 19 ± 2 per cent, by Weisskopf *et al.* (1978).

New proportional counters are under development which promise a great gain in the efficiency of polarisation measurements. These incorporate internal structure which gives signals dependent

Table 3.1 Spectral resolving power of detectors and dispersive spectrometers.

	Energy (keV)			
	0.5	1	6	20
Proportional counter RXTE	-	-	5.5	10
Imaging proportional counter ROSAT	2	3	-	-
CCD(ACIS) Chandra	5	11	40	-
CCD(MOS) XMM	8	14	38	-
Calorimeter	80	160	1000	
LETG Chandra	631	314	50	-
HETG/HEG Chandra	-	1351	225	-
RGS XMM	466	156	-	-

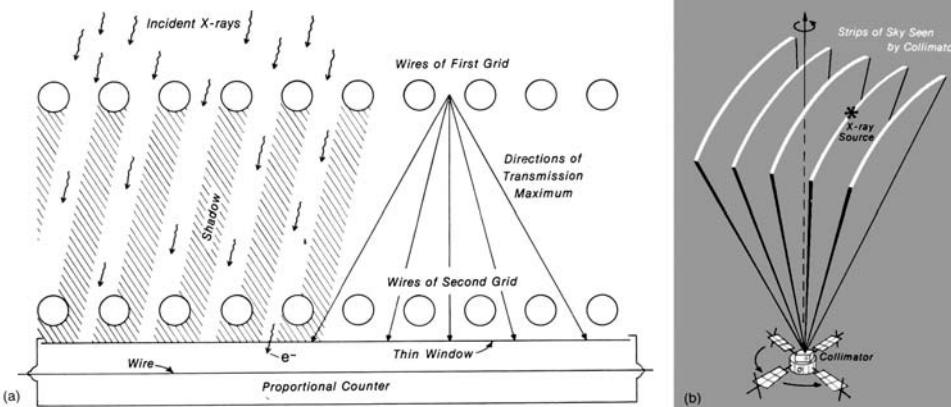


Fig 3.7 The modulation collimator. (a) Two separated wire grids define a set of bands on the sky through which X-rays can be seen. If the spacecraft rotates, a point source in the field of view will cross these bands and produce a modulated count rate as a function of time. (b) If the rotation axis is parallel to the pointing direction, the response pattern depends on where the source is within the overall field of view of the collimator, and the location of the source can be determined typically to $1'$.

on the direction of photoelectrons that are produced by X-ray interactions within the detector. Because photoelectrons are emitted in a direction approximately parallel to the electric vector of the incident X-ray, this directly measures polarisation. Two new designs are the gas pixel polarimeter (Bellazzini *et al.*, 2006) and the time projection chamber (Hill *et al.*, 2007).

3.2 Location of cosmic X-ray sources

3.2.1 Slat collimator

Early observations used detectors with simple slat or *honeycomb* collimators which restricted the detector field of view. One narrow and one broad dimension combined with a spinning spacecraft produced a scan of a strip of sky. Width of the strip was determined by the broad dimension. The narrow dimension produced a one-dimensional position of the source along the strip. A subsequent scan with different orientation would enable a two-dimensional location of sources on the sky with positional accuracy dependent on the collimator dimension and source strength. See figures in Chapter 1, e.g. Figs. 1.5 and 1.6

3.2.2 Scanning modulation collimator

In this scheme, a slat-collimated detector viewed the sky through a one-dimensional wire grid. As the detector scanned across the source, the signal was modulated by the shadow pattern of the grid, as shown in Fig. 3.7. The area of the detector was somewhat sacrificed to achieve this fine-scale modulation of the signal, which improved the location accuracy significantly. It essentially allowed the information about the spatial distribution of sources on the sky to be transformed into a temporal signal which could be easily registered by early X-ray astronomy detectors. The modulation collimator was used on a rocket flight in 1966 to locate Sco X-1 well enough to identify the optical counterpart (Gursky *et al.*, 1966). Arcminute locations of most of the bright sources were then obtained by modulation collimators on board the satellites SAS-3 and HEAO-1. Both linear and rotational scans were used.

3.2.3 Coded mask

An extension of the modulation collimator technique was developed in the 1970s, in which a *coded mask* is placed in front of a detector, making it possible to view a moderately large region of sky with a single pointing and with moderate spatial resolution. It is not an imaging device but more a pinhole camera with many pinholes. A sheet of absorbing material with a random pattern of

openings casts a shadow on a position-sensitive detector, as illustrated in Colour Plate 5. The pattern in the detector is correlated with the known pattern in the mask, and the two patterns match in only the direction to the source. If the mask is made larger than the detector, coverage of the field is uniform, except at the edges. Coded masks have been used for ‘imaging’ of high-energy X-rays and γ rays because they cannot be focussed in the same manner as low-energy X-rays (see Section 3.2.5). The data in Fig. 13.11 were obtained with an early-coded mask. Colour Plate 5 shows a schematic of a large coded detector on the INTEGRAL spacecraft.

3.2.4 Lunar occultation

A source can be accurately located in one dimension by observing it pass behind the limb of the Moon and exploiting our precise knowledge of the Moon’s position in the sky. As in the earlier sections, this converts temporal into spatial information. The orbital motion of the Moon is $\approx 0.5'' \text{ s}^{-1}$, and the resulting precision of the X-ray source location will depend on its strength. For an isolated source a rate of $\approx 10 \text{ counts s}^{-1}$ is needed for an accuracy of $1''$. The motion of the spacecraft must also be taken into account, so considerable preparation is needed to observe occultations. Figure 3.8 shows a lunar occultation of the Crab Nebula, a very bright X-ray source, observed with a large rocket-borne proportional counter. Colour Plate 9 shows an occultation of a galactic source observed by ROSAT.

3.2.5 Mirrors and telescopes

X-rays will reflect from smooth surfaces if the incidence angle is small. If the surface is truly smooth and the incidence angle $< 1^\circ$, reflection efficiency is close to 1. For a given energy, as the angle is increased, efficiency stays high until a critical angle, θ_c , after which reflection efficiency drops rapidly (as shown in Fig. 3.9 for some materials used in current-generation X-ray mirrors). For photon energy E and electron density ρ , then $\theta_c = c\sqrt{\rho}/E$. In a telescope, the angle is fixed. Reflection efficiency is high up to a point, then falls rapidly for energies such that the reflection angle is greater than the critical angle. Thus telescope mirrors have a high energy cut-off,

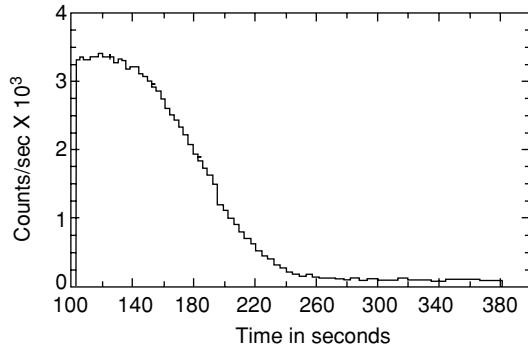


Fig 3.8 Counting rate of a rocket-borne X-ray detector during a lunar occultation of the Crab Nebula. As the Moon passed in front of the nebula, the source within the Crab was gradually occulted over 120 s of time, showing that it is $\approx 2'$ in extent. The neutron star within the Crab, however, accounts for 5 per cent of the X-ray flux and is a point source. When the neutron star passed behind the lunar limb at 195 s, a sudden drop in count rate was observed (from Palmieri et al., 1975).

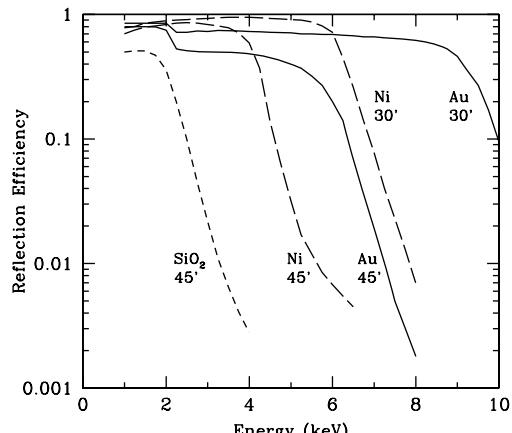


Fig 3.9 Efficiency of X-ray reflection from polished surfaces of quartz (short-dashed line), nickel (long-dashed line) and gold (solid line). Incidence angles are 0.5° (30 arcmin) and 0.75° (45 arcmin). Reflectivity is highest at low energies and small angles (from Seward, 1977).

and dense (high-Z) materials have the largest critical angles. At 0.75° X-rays are reflected from polished quartz up to $\approx 2 \text{ keV}$, from polished nickel up to $\approx 4 \text{ keV}$ and from polished gold up to $\approx 6 \text{ keV}$. At 0.5° the Ni and Au cutoffs are raised to 6 and 9 keV, respectively.

A parabolic mirror surface can be made or approximated which, with one reflection, will

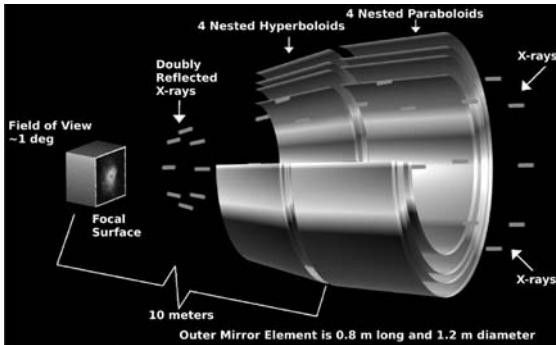


Fig 3.10 The geometry of the Chandra mirrors. Reflection angles are greatly exaggerated. An annular on-axis beam enters the aperture of each mirror pair. X-rays are first reflected from the inner surface of the parabolic mirror and then from the inner surface of the hyperbolic mirror. The beam comes to a focus at the detector (NASA/CXC/D. Berry 2010).

focus a parallel beam of X-rays to a point. Such telescopes have been used to concentrate X-rays onto a small spectrometer or polarimeter. Two reflections, however, are required to produce an image. Imaging geometries were studied by Wolter (1952) for X-ray microscopy. The design usually adopted for a telescope is the Wolter type 1, where X-rays are first reflected from the inside of a parabolic surface of revolution and the second reflection is from the inside of a hyperbolic surface. This is shown in Fig. 3.10, which also illustrates the nesting of mirrors to increase the telescope's effective area.

The on-axis image of a point X-ray source produced by such a telescope consists of a bright core surrounded by a faint, broad halo. The size of the core is determined by the figure of the mirrors, while the broad halo arises from scattering by small imperfections in the mirrors. The surface brightness distribution of the image of a point source is called the *point spread function* (PSF). A mirror is also characterised by the *half-power radius*, the radius of a circle containing half the power of the incident beam. Figure 3.11 shows the dependence of the Chandra half-power radius on off-axis angle and energy. Off-axis, the beam spreads and the image becomes asymmetrical. At higher energies, there is more scattering. Although the Chandra field of view is quoted as 16' or 40' (for the ACIS-I and HRC-I detectors), only the

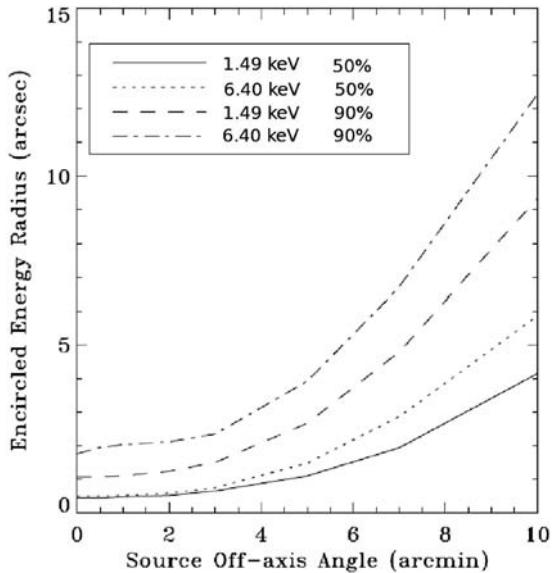


Fig 3.11 Size of Chandra focal spot for two energies. The radii which include 50 per cent and 90 per cent of the power are shown as a function of off-axis angle. Note that the spot is larger for larger energy and that the region of best focus is a field of radius $\approx 2'$ (CXC 2007).

central 2' contains the best-quality image. Colour Plates 18 and 54 show this characteristic of X-ray telescopes.

Figure 3.6 shows an on-axis Chandra image of a bright source and a projection of this image in one dimension. The core is overexposed so as to show the broad scattering wings of the image, which are clearly visible in the projection. The projection also illustrates the advantage of an imaging telescope over the old, mechanical collimators.

For most X-ray detectors the major background contribution to the counting rate comes from cosmic rays and solar particles. This background is proportional to the detector area, and to be detectable, the signal from a source must be significantly greater than the background. The telescope, however, focusses the X-rays onto a small area of the detector, and only the background from this area must be overcome. The gain of the Einstein imaging system over the HEAO-1 scans was about a factor of 1000 in source detection. The Chandra image of a point source is so small that there are no background events in an exposure of a few hours. The projection in Fig. 3.6 shows the

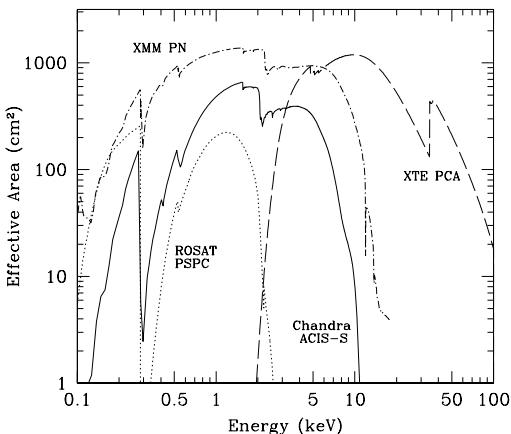


Fig 3.12 The effective area of the principal instruments on several X-ray missions: proportional counters on ROSAT and XTE; CCD detectors on Chandra and XMM. Note that there is almost no overlap in the energy ranges covered by ROSAT and XTE (PIMMS, 2008).

equivalent of a slat-collimator scan over this small region. Note the high contribution of the background in the scan across the two weak sources (and this background is only from a few columns of the CCD detector). It is the tight focus of an imaging system which gives the ability to detect very faint sources.

3.3 | Spectroscopy

3.3.1 Spectral analysis, detector efficiency

Detector efficiency is expressed as *effective area*. This is the window area minus any shadowing by the collimator, multiplied by the probability of transmission through the window material, multiplied by the probability of absorption in the active volume of the detector. If a telescope mirror is used, the efficiency of reflection from the mirror is included. Transmission, absorption and reflection all vary with energy. Figure 3.12 shows effective areas for some modern systems. Generally, the window transmission limits the low-energy response, and the high-energy cutoff is set by the mirror (Fig. 3.9) and/or by the detector becoming transparent.

Most spectra, to date, have been obtained with proportional counters and CCD arrays, which have modest energy resolution. With these data the

intrinsic source spectrum is inferred by running a trial spectrum through the detector on the computer and then comparing the predicted output with the measured pulse height spectrum. The trial spectrum is first multiplied by the effective area then convolved with the detector resolution and compared with the data. Spectral parameters are adjusted to get the best fit. It is customary to include absorption in the ISM, N_H , in the trial spectrum and to quote an unabsorbed flux (calculated with N_H set to 0) which can be used to calculate the luminosity of the source.

Box 3.2 Spectral resolution and resolving power

An event in a proportional counter or CCD produces a number of ions N , which is proportional to photon energy E . At any given energy, the signal from the detector is broadened because of statistical fluctuations in the number of ions. This broadening, ΔE , is usually taken as the full width half maximum (FWHM) of the response to monoenergetic photons.

The precision with which the energy of an incident photon can be determined is described by the *spectral resolving power*, $R = E/\Delta E$, or by the *energy resolution*, a term which does not have a precise definition. For example, a proportional counter is said to have a resolution of 15 per cent at 6 keV ($\Delta E/E$), whereas a calorimeter is said to have a resolution of 10 eV (ΔE).

For a given E , and N not too small, N is governed by Poisson statistics, and so its standard deviation is \sqrt{N} . In these detectors, $E/\Delta E \propto N/\Delta N \propto \sqrt{N}$, and R increases with energy.

On the other hand, the resolution of an X-ray diffraction grating is set by the spatial resolution of the telescope, i.e. the spread in size of the image is translated to a spread in wavelength $\Delta\lambda$, which is approximately constant over the dispersed spectrum. Thus $R = E/\Delta E = \lambda/\Delta\lambda$, which is now $\propto \lambda \propto 1/E$, and R decreases as energy increases.

3.3.2 Diffractive spectroscopy

Diffraction gratings with high resolving power can be used to obtain detailed X-ray spectra.

Table 3.2 | X-ray telescope mirrors.

Mission	Aperture diam. ^a (cm)	Number nested mirrors*	Geom. area [*] (cm ²)	Grazing angles (')	Focal length (m)	Mirror coating	Max. energy (keV)	On-axis resolution (")
Einstein	58	4	350	40–70	3.45	Ni	4.5	3
EXOSAT	28	2	80	90–110	1.09	Au	18	2.4
ROSAT	83	4	1140	83–135	2.4	Au	2.4	3
ASCA	40 (1)	120 (1)	1300 (2)	21–45	3.8	Au	12	120
Chandra	120	4	1500	27–51	10	Ir	10	0.5
XMM	70 (1)	58 (1)	4650 (3)	18–40	7.5	Ir, Au	12	6
Suzaku	40 (1)	175 (1)	3480 (4)	11–36	4.7	Au	15	60

^a Number of modules in parentheses.

At energies below 1 keV the resolving power is 100 times greater than that of CCD detectors. In the 1960s rocket-borne instruments obtained grating X-ray spectra of the brightest source in the sky, namely, the sun (e.g. Evans & Pounds, 1968). In the period \sim 1965–1975, spectrometers on a series of OSO spacecraft routinely measured solar X-ray lines with high spectral resolution (e.g. Meekins *et al.*, 1970). Now the development of large X-ray mirrors (see Table 3.2), particularly Chandra and XMM, has made the use of gratings possible in observations of hundreds of other cosmic sources.

Chandra carries three sets of transmission gratings, which can be inserted into the beam of the telescope just behind the mirrors. The fabrication of these gratings was a challenge. The wavelength of a 1-keV X-ray is 12.38 \AA . To have a 1-keV line displaced by 1 cm over the 10-m focal length requires constructive interference at an angle of $\approx 10^{-3}$ radians or $\approx 3'$. The grating ruling spacing p must satisfy $\sin\beta = \lambda/p$ so $p \approx 10^{-7}/10^{-3} = 10^{-4} \text{ cm} = 1 \mu\text{m}$. The three sets of Chandra gratings have spacings of 0.2, 0.4 and 1.0 μm and are shown schematically in Fig. 3.13. Because the grating bars must be opaque, they are made from gold, a high-Z material, and the high-energy gratings are 2.5 times thicker than the grating ruling spacing – a difficult fabrication. There is some transmission through the gold, so bar width and thickness were set to maximise the first-order diffracted intensity (at 1.5 keV, 38% of the energy goes into the first order [both sides] and 8% into the zeroth order).

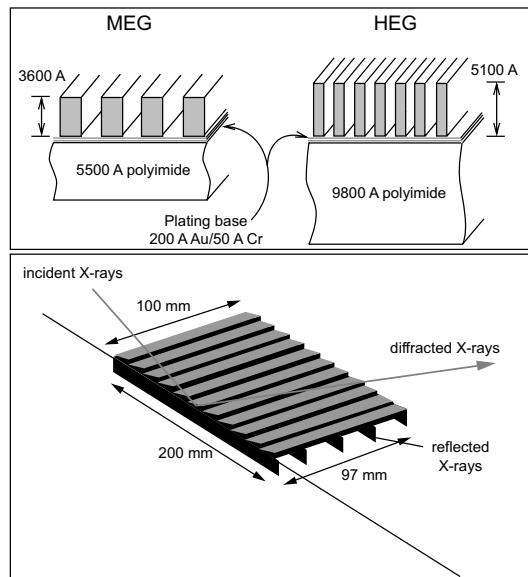


Fig 3.13 Schematic of Chandra and XMM diffraction gratings. The Chandra Medium Energy (MEG) and High Energy (HEG) gratings are transmission gratings with 2500 and 5000 lines/mm, respectively. The gold bars are thicker than the grating openings, and a thin plastic support is needed for strength. The XMM gratings are reflection gratings with 645 lines/mm. The incident angle is 1.5° . (NASA/MIT 2000; ESA 2009).

The geometry, of course, is not perfect. There are irregularities in the grating structure and corners are rounded, so a careful preflight calibration was required. Chandra spectra are shown in Figs. 8.20, 6.17 and Colour Plate 13. For details, see Canizares *et al.* (2005) and CXC (2007).

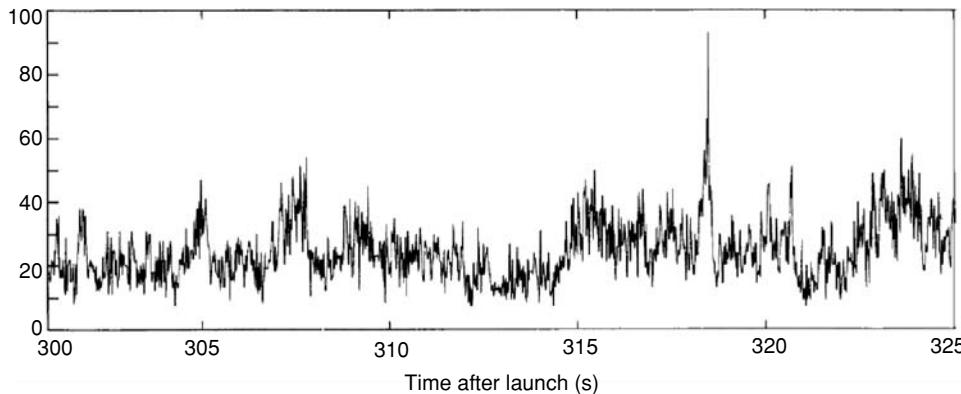


Fig 3.14 Count rate of a rocket-borne proportional counter pointed at Cyg X-1 in October 1973. The y axis shows counts per 20 ms, so the typical count rate was about 1000 counts s⁻¹. Note the strong, irregular and rapid variation of X-ray emission (adapted from Rothschild et al., 1974).

XMM carries a set of reflection gratings on two of the three XMM telescopes. These are permanently mounted in the beams, aft of the nested mirrors. Half the energy is intercepted and diffracted to two CCD arrays devoted to the spectrometers. Thus spectra in the range 0.3–2.5 keV are obtained for every source observed. The gratings are reflection gratings, and the X-rays are incident at an angle of 1.5°. Again, fabrication was a challenge. Grating elements have SiC substrates, are coated with gold and have a ruling spacing of 1.5 μ. XMM spectra are shown in Figs. 4.10 and 13.9. For details, see XMM_RGS (2008).

Spectra from lower-resolution detectors, such as CCDs, are shaped by the X-ray source's continuum. Lines are only clear in spectra which have a few thousand counts or more. In contrast to this, spectra obtained with diffraction gratings show lines clearly. Emission lines are prominent, but the continuum can be faint and subject to background uncertainty. Indeed, one of the most interesting aspects of grating spectra is the line shape. Velocities and absorption features can be recognised in the highest-resolution spectra.

3.4 Timing

The time history of emission from a source conveys much information and often is clearly sufficient to identify the nature of the source. These

data form part of the narrative in the chapters which follow.

Even early rocket flights showed that many of the bright sources were variable and that the variability could be irregular and rapid (Fig. 3.14). These were the accretion-powered binaries. Accretion-powered active galactic nucleus (AGN) also varied irregularly but on a longer timescale (Fig. 14.29). Coronal emission from stars varies unevenly and shows flares, as in the Sun, which last minutes to hours. The rotating neutron stars exhibit extremely regular pulsations, some almost sinusoidal, some sharply peaked (Fig. 9.2). Light curves also show long-term variations and eclipses (Fig. 11.2). Technically, the large diffuse sources also vary, but not on the timescales of our observations or even our lifetimes. At the other end of the scale, outbursts of emission define the nature of supernovae, γ-ray bursters and transient X-ray sources.

When data are obtained from a new source, one of the first things examined is the light curve. If there are a few hundred events or more, a search for pulsations can be attempted. A fast Fourier transform (FFT) is applied, and promising frequencies are further investigated with epoch folding – data are summed modulo a range of frequencies centred on those with high power identified by the FFT. If the emission is periodic, these data yield the pulse shape. The more sensitive FFT programs search for power over several summed harmonics

and allow for variation of frequency over the time of the observation. Such variation is common and arises from Doppler shifts in binary systems or pulsar spindown. Pulsed emission from bright sources can be obvious in the initial light curves, but detection of regular pulsations close to threshold is a real art. The measurement of quasiperiodic oscillations from accretion-powered binaries is also difficult and is covered in Chapter 11.

3.5 | Significant missions

3.5.1 Uhuru

The first satellite devoted to X-ray astronomy, Uhuru, was launched into a near-Earth equatorial orbit from a platform off the coast of Kenya. The payload, at 64 kg, weighed no more than a typical rocket experiment at the time. It carried two sets of conventional proportional counters with simple honeycomb collimators, which were used to undertake the first X-ray survey. The spacecraft was spin stabilised at 12 min revolution⁻¹. One detector had a field of view of $1^\circ \times 10^\circ$, so it viewed each source for 2 s during each scan. Uhuru scanned many times (typically, ~ 60) over the same region, thereby greatly increasing its sensitivity to weak sources. The net result of this was that Uhuru was able to detect X-ray sources 10 times fainter than the faintest detectable on earlier rocket flights. Although not uniform in sensitivity, 95 per cent of the sky was scanned during the 2.5-year lifetime of the mission. A tape recorder failed early on, and the transmitter failed and then miraculously recovered. Under these circumstances, efficiency was not optimum. Nevertheless, the 4U (fourth Uhuru) catalogue contains 339 sources. Strong sources were located with 2' accuracy, whereas weak sources could only be placed within a region covering several square degrees. Spectacular light curves of the brightest sources proved that some were binary systems, and spin and orbital periods were measured. Details are given by Giacconi *et al.* (1971) and Forman *et al.* (1978).

3.5.2 EXOSAT: A European observatory

EXOSAT was an European Space Agency mission: a collaboration of England, Germany, Italy and

the Netherlands. Originally, it was designed to observe lunar occultations, and so it was launched into a highly elliptical orbit to increase sky coverage. However, by the time of the 1986 launch, the brighter sources had already been well located by other missions, so observations of occultations were not attempted. The spacecraft carried two small telescopes capable of 20" resolution and a large proportional counter array for timing and variability studies. The elliptical orbit extended to a distance of 30 Earth radii and took 4 days to complete one revolution. Only during perigee (i.e. when passing closest to the Earth) was the satellite inoperative as it rapidly crossed the Earth's radiation belts. The orbit allowed up to 76 hours of continuous viewing for most sources, a unique asset at the time. Indeed, the major accomplishments of the mission were measurements of high-quality, long-duration light curves and pulsations from accretion-powered sources. Details are in White and Peacock (1988) and in Chapter 11.

3.5.3 HEAO-1: An all-sky survey

Launched August 1977, the first High Energy Astronomy Observatory carried four large instruments all with mechanical collimators. An all-sky survey was accomplished, then the instruments were pointed at various sources to study spectra and time variability. The principal survey instrument was the A1 proportional counter array sensitive to energies of 1–20 keV. This produced a catalogue of 842 sources (see Colour Plate 2). Weaker sources were located to 0.5° accuracy and the stronger ones to 1'. The A2 detectors covered the energy range 0.2–60 keV. These were designed to measure spectra of sources and in particular of the diffuse background (see Fig. 16.2). The A3 system was a bank of proportional counters behind modulation collimators. Moderately bright sources could be located with an accuracy of 20", which was generally good enough to identify the optical counterpart. The A4 system was a set of phoswich scintillation detectors to detect hard X-rays and γ rays in the 15-keV to 10-MeV range. The sky survey found 70 sources in the range 13–180 keV. Details are in Bradt *et al.* (1992).

3.5.4 Einstein

Launched in November 1978, HEAO-B (later to become known as the Einstein Observatory) carried the first serious X-ray telescope. A set of 4 Wolter type 1 nested mirrors focussed X-rays up to 8 keV in energy. Spatial resolution was 5'' on the axis, degrading to 1.5' at the edge of the 1° field of view.

Einstein employed a turntable by which different detectors could be placed at the telescope focus. X-ray imaging was the principal aim of the mission. Both an IPC and micro-channel plate detector (HRI) were included. The HRI's spatial resolution was that of the mirrors, but energy resolution was poor. The IPC recorded the energy of each event, but spatial resolution was $\approx 40''$ at best. A third detector, the (SSS), provided a factor of ≈ 3 better energy resolution but no spatial information because it recorded all events occurring in its 6' field of view. A monitoring proportional counter (MPC) was coaligned with the telescope to monitor variability of the brighter X-ray sources.

Two high-resolution spectrometers were included. The objective grating spectrometer (OGS) was a set of gold diffraction gratings which could be inserted into the telescope light path. The dispersed spectrum was recorded by one of the imagers. The focal-plane crystal spectrometer (FPCS) was a curved-crystal Bragg spectrometer which operated at the telescope focus. Effective area of these spectrometers was very small ($\approx 1 \text{ cm}^2$ or less), so they were used only on the brightest sources.

The SSS had a lifetime of 11 months because of a limited onboard cryogen supply. The observatory itself lost its ability to point after only 2.5 years in orbit. Operations ceased after a series of gyro problems forced the rapid use of all the attitude-control gas. Einstein accomplished 5600 observations. For the first time we saw the structure of supernova remnants, the distribution of sources in normal galaxies and the structure of clusters of galaxies. Deep observations of blank fields resolved 25 per cent of the diffuse background into discrete sources. Images from Einstein can be seen in Figs. 5.6 and 7.1. Observatory details are in Giacconi *et al.* (1979)

3.5.5 Granat: High-energy images

Granat (meaning 'pomegranate', a fruit with several chambers and many seeds) was a Russian spacecraft with the participation of France and Denmark. It was placed in a highly elliptical orbit (4-day period) on 1 December 1989 to search for hard X-ray sources and γ -ray bursts. It carried two large X-ray telescopes which viewed a small region of the sky through coded masks, an X-ray proportional counter, an all-sky monitor and γ -ray burst detectors. The coded-mask imaging technique (described in Section 5.2.3) was a radical departure from the techniques of ground-based astronomy.

The first telescope, Sigma, mapped selected regions in the 35–1300 keV range. The second telescope, ART-P, was a four-module system sensitive from 4–100 keV and co-aligned with Sigma. The ART-P energy range extended from that of conventional X-ray astronomy to the unknown territory of higher energies. The two telescopes together made maps of the high-energy sky, with particular emphasis on the galactic centre. The sky was also monitored for γ -ray bursts.

The higher energies are important. This is where the more extreme characteristics of sources appear. The highest energies are usually generated the closest to the central engines of AGN and may carry the clearest signature of their exotic properties. This is also true for accretion-powered compact sources which contain neutron stars and stellar-sized black holes. Also, above 5 keV, absorption in the galactic plane is minimal, meaning that a high-energy survey of our Galaxy should reveal sources in spiral arms on the far side of the galactic nucleus. Granat details are in Sunyaev *et al.* (1990). Accomplishments include monitoring of sources around the galactic centre and γ -ray burst afterglow observations (see Chapter 17). Figure 13.11 shows a Sigma image of the galactic centre.

3.5.6 ROSAT: Deep all-sky survey

The ROSAT, a German mission with U.S. and U.K. participation, was launched 1 June 1990. The principal instrument, an X-ray telescope, was built for high sensitivity at low photon energies (< 2 keV). Four nested reflectors focussed X-rays on either

of three imaging detectors. The effective area for X-ray collection was twice that of Einstein, but it had no sensitivity above 2.4 keV. With larger angles of incidence on the mirrors, it was easier to achieve a large collecting area, but the high-energy response was sacrificed (see Fig. 3.12).

The primary detectors were two (for redundancy) gas-filled position-sensitive proportional counters, the PSPCs. A PSPC, when at the focus of the telescope, had a field of view of 2° . The first 6 months of the mission were used to conduct an all-sky survey with a PSPC at the focus. One-half of the time produced useful data (all that could be expected with the near-Earth orbit). The scan was perpendicular to the plane of the orbit, so coverage of the sky was not uniform, with celestial objects in the orbital plane viewed for ≈ 600 s, whereas those close to the pole of the orbit were viewed for ≈ 4000 s. Colour Plate 3 shows the ROSAT view of the soft X-ray sky which resulted from this survey. The catalogues of Voges *et al.* (1999, 2000) list more than 18 000 bright sources and more than 105 000 faint sources.

The second focal plane detector was the HRI (provided by the United States), similar to that carried by Einstein but more efficient. There was also a separate wide field camera supplied by the United Kingdom. This was designed to point parallel to the X-ray telescope and to map the sky in the extreme UV (60–300 Å). It incorporated a mirror with mean grazing angle 7.5° , geometrical area 475 cm^2 and an HRI-like detector. A total of 479 sources were found, mostly late-type coronal stars and hot white dwarfs.

A 580-k circular orbit at 56° inclination allowed the satellite to pass over a ground station at Weilheim, Germany (latitude 48°), where all data were received. ROSAT re-entered the atmosphere in early 1999 after 9 years of operation. The survey was 90 per cent accomplished in the first 6 months. A brief loss of control caused the satellite to tumble, and the first PSPC was lost when it pointed briefly at the Sun. The survey was completed using the second PSPC, and at other times the telescope was used for pointed operations. The PSPC was turned off when the gas supply was almost exhausted, and observations continued with the HRI until a few days before re-entry. At this point the PSPC was refilled

and used for observations until atmospheric drag rendered the spacecraft uncontrollable. ROSAT produced soft X-ray maps of large structures that are unsurpassed. See e.g. Figs. 8.22, 9.11 and Colour Plate 56. Deep observations resolved 75 per cent of the soft diffuse background, and X-ray emission was discovered from comets and the Moon. Observatory details are in Aschenbach (1988) and Pfeffermann *et al.* (1987).

3.5.7 ASCA

The Japanese spacecraft Astro-D, after launch in early 1993, was renamed ‘Asuka’ (flying bird), which was then shortened to ASCA (for Advanced Satellite for Cosmology and Astrophysics). A novel method was used to make the telescope, with concentric assemblies of mass-produced thin-foil conical mirrors as reflectors (Serlemitsos *et al.*, 1995). Spatial resolution was relatively poor (only $3'$), but collecting area was large, and mirrors were fabricated easily and cheaply. Because the mirrors were thin, the shadow cast by each mirror element was minimal. There were four mirror assemblies, each with 120 concentric mirrors. Two focussed on CCD arrays (the SIS) and two on imaging gas scintillation proportional counters (IGSPC or the GIS).

Both detector systems were used for imaging spectroscopy and timing. The two systems were complementary. The CCD arrays had superior spectral resolving power but no timing capability below 0.1 s. They were also susceptible to radiation damage from intense particle bombardment in the Earth’s radiation belts.

The IGSPC had a larger field of view, was not as vulnerable to radiation damage and was capable of measuring event times to 0.1 ms. Spectral resolution, however, was a factor of 4 worse than that of the CCD detector. Even so, the IGSPC spectral resolution was a factor of 2 better than conventional proportional counters (as used e.g. on ROSAT). Tanaka *et al.* (1994) give a description of the observatory.

3.5.8 Rossi XTE: High temporal resolution

The X-ray Timing Explorer (XTE) mission, launched at the very end of 1995 (and still operating 13 years later), was designed for fast X-ray timing. It monitors, with millisecond accuracy, the emission of strong and moderately strong X-ray sources which

are reasonably well separated from their neighbours. It is also equipped to detect and monitor transient sources over the whole sky. RXTE carries three sets of detectors: a large-area proportional counter a large-area scintillation counter and an all-sky monitor.

The proportional counter array (PCA) has a large collecting area and is the most sensitive instrument on the payload. It consists of five identical methane-xenon-filled proportional counters. The area has been maximised to collect as many photons as possible in very short time intervals. A 1-m-long honeycomb collimator restricts the field of view to 1° FWHM. Each detector has dimensions of approximately $0.3 \times 1.0 \times 1.3$ m, and the total weight of the PCA detectors alone is three-quarters of a tonne. This ‘small’ Explorer-class mission is actually not very small physically! The High-Energy X-ray Timing Experiment, HEXTE, is mounted alongside the PCA, pointed in the same direction and with the same field of view. The detectors are caesium iodide/sodium iodide scintillators: eight crystals viewed by eight photomultiplier tubes, giving a large effective area and covering energies up to 200 keV.

The PCA and HEXTE do long-term monitoring of X-ray luminosity and enable broadband spectral monitoring of bright X-ray sources with high time resolution. The time resolution, of course, depends on the source strength. The satellite is capable of measuring photon arrival time to an accuracy of 5 μ s. It can observe millisecond fluctuations from stronger sources and variations with durations of hours to days from the fainter ones.

The third detector is an all-sky monitor (ASM) which uses instruments with the wonderful name of ‘X-ray shadow cameras’. One-dimensional coded masks cast shadows of the X-ray sky on small position-sensitive proportional counters. Each source produces a shadow with a distinctive pattern. A computer decodes the data, unscrambles the overlapping shadows and derives the location of each source. By rotating these detectors it is possible to scan the whole sky every 1.5 hours. This continuously monitors about 50 strong sources in our Galaxy and has produced a record of daily flux intensity spanning more than a decade for each source. The sky is also surveyed

for transient sources, such as X-ray novae, which appear unpredictably; details can be obtained from http://xte.mit.edu/ASM_lc.html.

The spacecraft can point at any location in the sky, except within a small region close to the Sun. It can slew rapidly within 1 hour from one target to another. The principal instruments can be pointed at a transient source within hours of discovery and certainly within a day or two of the outburst. The instruments can also be pointed in the antisolar direction, a capability desirable for simultaneous optical and X-ray observations. Ground-based observatories are, of course, constrained to looking during the night. Some notable results have been timing studies of binaries, including discovery of quasi-periodic oscillations; observation of an eclipse in η Car; and monitoring of outbursts from Soft Gamma-ray Repeaters, Anomalous X-ray Pulsars, and transient sources. Instrument details are in Glasser *et al.* (1994), Gruber (1996) and Levine *et al.* (1996).

3.5.9 BeppoSAX: Broadband spectroscopy

This was a major Italian program with participation of the Netherlands. Several instruments of modest size performed observations over a broad energy band of 0.1–200 keV. Instruments were coaligned and pointed with an accuracy of $2'$. The purpose of the mission was to determine source properties through timing and spectroscopy.

The broad energy coverage was designed for accurate determination of spectral parameters. To collect photons below 10 keV, there were four nested, gold-coated mirror/concentrator assemblies focussing on position-sensitive gas scintillation proportional counters. A high-energy gas scintillation proportional counter, the HEGSPC, extended coverage to 120 keV, and a phoswich scintillation detector, the PDS, recorded photons with energies as high as 200 keV.

For long-term monitoring of strong sources and detection of transients, two wide-field cameras, the WFCs, completed the payload. Each utilised a xenon-filled position-sensitive proportional counter (PSPC) behind a coded mask. BeppoSAX was launched in April 1996 and re-entered the atmosphere almost exactly 7 years later. It accomplished the first observation of a γ -ray

burst X-ray afterglow and monitored the afterglow intensity of many γ -ray burst sources. Broadband spectroscopy was obtained for all types of X-ray sources. For details, see Boella *et al.* (1997).

3.5.10 Chandra: Really high spatial resolution

Chandra (née AXAF, the Advanced X-ray Astrophysics Facility) was designed to have 4 times the area of the Einstein mirror at low energies and to have a considerable collecting area between 6 and 7 keV, the energy of the iron lines emitted by most astrophysical sources. An order of magnitude better angular resolution was achieved, and imaging sensitivity is 2 orders of magnitude better than that of Einstein.

The telescope consists of a set of eight nested reflecting surfaces, all arranged in Wolter type 1 geometry (Fig. 3.10). The high-energy response is achieved by use of small reflection angles and by coating the three inner mirrors with iridium. Angular resolution is a stunning 0.5 arcsec. The combination of high resolution, large collecting area and sensitivity to higher-energy X-rays makes possible the study of extremely faint sources in crowded fields.

Chandra was placed in near-Earth orbit from the Shuttle Columbia in July 1999. An attached engine propelled the spacecraft to the final orbit which has a high eccentricity and varies somewhat with time. Apogee is at 20 Earth radii and perigee is at 3 Earth radii. This 63.5-hour orbit is more eccentric than that shown in Fig. 1.2 and places the observatory outside the magnetosphere much of the time. Seventy per cent of the time is spent above the radiation belts, and continuous observations up to 160 ks long are possible. However, if the Sun is active and there is a high flux of solar particles, the instruments are shut down to avoid damage. A few days observing each year are lost because of solar activity. Overall, observing efficiency has been 60 to 70 per cent. The solar aspect angle is limited to between 45° and 170° , so 85 per cent of the sky is accessible at any time, which allows Chandra to respond to about 40 targets of opportunity (TOO) requests every year.

There are four focal plane instruments and a mechanism to place any one of them at the telescope focus. The Advanced CCD Imaging

Box 3.3 | The Chandra mirror array

The mirror array was designed and nurtured by Leon Van Speybroeck, the ‘Telescope Scientist’ (Fig. 3.15). Years were spent in planning and design. He supervised the grinding, polishing, alignment and testing of the mirrors. During horizontal testing of the finished mirror elements, he found that gravity was distorting the thin Zerodur glass shells and worked out a method for compensating for this effect. Performance of the mirrors in orbit has been superb. The very first time the mirror-detector combination was exposed to space (the ‘moment of truth’), X-rays from a weak serendipitous source in the field started to accumulate in a small region of the detector. Seeing the image of this source, it was immediately clear that all was well and performance would be as expected. The source was promptly named Leon X-1 (aka CXOU J051936.3-604804, an AGN at $z = 0.32$).



Fig 3.15 The Chandra ‘Telescope Scientist’, Leon Van Speybroeck, in his office 2 years before launch. At this time the mirror elements had been assembled and tested. The ground calibration with X-rays at MSFC Huntsville had been completed, and the telescope was known to be capable of sub-arcsecond resolution. Only final spacecraft assembly and launch remained before observations could begin.

Spectrometer (ACIS) is a collection of 10 CCD chips, four in an imaging array (which has a $16' \times 16'$ field of view) and six in a linear array for spectroscopy (see Fig. 3.5). The HRC also comprises two detectors: HRC-I is a square micro-channel plate with a field of view of $40'$; HRC-S was made for soft X-ray spectroscopy and has a $6' \times 100'$ field of view that covers the low-energy transmission grating (LETG) dispersion pattern from 0.07 to 10 keV. The HRC detectors were included to accomplish high spatial resolution and fast timing. However, a wiring error in HRC-I reduced its timing capability, so HRC-S is used instead for Chandra fast-timing observations.

Three diffraction gratings can be rotated into the telescope beam, and the diffracted spectra are imaged with HRC-S and ACIS-S detectors. For the important features of the Chandra spacecraft, see Colour Plate 6. The telescope-grating combinations are sensitive enough to record high-resolution spectra from many types of moderately bright sources, and many examples are included in this book. More details are in Weisskopf *et al.* (2002), CXC (2007), at <http://cxc.harvard.edu/> and at <http://chandra.harvard.edu/>.

3.5.11 XMM-Newton: High-throughput spectroscopy

The European X-ray Multi-Mirror Mission was launched in late 1999. Like Chandra, it is designed for a combination of imaging spectroscopy and high-resolution grating spectroscopy. Using more mirrors gives an effective area of the imaging spectrometers 4 times that of Chandra, but the focus is not as sharp. The grating spectrometers have twice the Chandra area but half the spectral resolving power. These two X-ray observatories, operating at the same time, are complementary to each other.

XMM, too, is in an elliptical orbit with a 48-hour period, which extends from 1 to 18 Earth radii and allows observations as long as 130 ks between passes through the radiation belts. Power requirements limit the pointing direction to solar aspect angles between 70° and 110° . Like ROSAT, most sources can only be observed during two intervals each year. The chance of being able to view a fast TOO is about 30 per cent.

The spacecraft carries one optical/UV and three X-ray telescopes. Each X-ray telescope is made from

58 nested Wolter type 1 reflectors. The reflectors are gold deposited on thin electroformed nickel elements, each being a parabolic-hyperbolic pair. Each X-ray telescope has a CCD at the focal plane, all provided by the European Photon Imaging Camera Consortium (EPIC). Two telescopes with EPIC metal-oxide semiconductor (MOS) CCDs also have reflection gratings (RGS) which intercept 58 per cent of the beam, producing dispersed spectra which are recorded on separate CCD arrays. The third telescope has EPIC PN CCDs used solely for spectral imaging. There is no switching of instruments. All detectors operate for each source, but a variety of operational modes for the CCDs are possible. Because the focus is not as tight, the pileup problem is not as severe as that of Chandra. Trouble-free count rates can be 10–100 times higher with the EPIC detectors.

The RGS spectrometers record data for every observation over the energy range 0.34 to 2.5 keV and cover both first and second order. Figure 13.9 shows an RGS spectrum of a diffuse source. EPIC images can be found e.g. in Colour Plates 27 and 48. Also see Jansen *et al.* (2001) and XMM_UH (2008).

3.5.12 INTEGRAL: A γ -ray mission

The International Gamma Ray Astrophysics Laboratory, launched in October 2002, is a European γ observatory with coverage extending into the X-ray band. The spacecraft carries three instruments which image via coded masks: the SPI γ -ray spectrometer covers the range 18 keV to 8 MeV, viewing a 16° patch of sky with an angular resolution of 2.5° . Detectors are high-purity germanium and give a spectral resolution of 2 keV. The IBIS imager (see Colour Plate 5) covers the range 15 keV to 10 MeV and views the same region but with angular resolution of 0.2° . The JEM-X system provides images with arcminute angular resolution in the 3–35 keV energy band. Thus both high-resolution spectra and accurate locations are obtained for e.g. transient γ -ray sources.

3.5.13 Swift: Fast response for γ -ray bursts

The Swift spacecraft, launched in November 2004, is dedicated to the study of γ -ray bursts (see

Box 3.4 | Characteristics of detectors on some observatories

Not all missions are listed nor all detectors for each mission (Table 3.3). More complete lists can be found in Bradt *et al.* (1992) and Seward (2000). Effective areas vary over the detector energy range. The area given is the approximate maximum area and usually applies to a broad band in the middle of the energy range. If slat collimator, the field of

view is FWHM of the collimator response. Effective areas and Crab rates refer to all units of the system listed e.g. the sum of both XMM EPIC/MOS telescopes. Crab rates are from Seward (2000) and PIMMS (2008). Grating areas and rates apply to the first-order diffracted spectrum only.

Table 3.3 Selected missions and detector characteristics.

Mission	Dates (launch–demise)	Detector	Energy range (keV)	Effective area (cm ²)	Field of view	Spatial resolution	Crab signal (cts s ⁻¹)
Uhuru	70–73	PC1	2–18	840	1° × 10°	2'–1°	950
		PC2	2–18	840	10°	–	
Exosat	83–86	ME	1–35	1600	45'	–	1740
		CMA	0.1–2	10	2°	20"	
HEAO 1	77–79	A1	1–20	3300	1° × 4°	–	13 600
		A2	0.5–3	800	1.5° × 3°	–	
		A4	15–200	200	1.7° × 20°	–	
Einstein	79–81	IPC	0.2–4.5	150	1.2°	40"	680
		HRI	0.2–4	15	25'	5"	
		SSS	0.4–4	180	6'	–	
		MPC	1.2–20	670	45'	–	
		OGS	0.2–2	0.6	–	5"	
		FPCS	0.2–3.4	1	–	–	
Granat	89–98	sigma	35–1300	800	5°	13'	30
		ART-P	4–60	1250	1.8°	5'	
ROSAT	90–99	PSPC	0.2–2	250	2°	30"	960
		HRI	0.2–2	80	40'	2"	
		WFC	0.04–0.2	10	5°	3'	
ASCA	93–01	SIS	0.4–10	200	24'	3'	1020
		GIS	0.7–10	150	24'	3'	
RXTE	95–?	PCA	1–20	7400	1°	1°	11 800
		HEXTE	12–60	1000	1°	1°	
BeppoSAX	96–03	MECS	1.3–10	450	0.5°	1.2'	960
		HPGSPC	4–120	260	1.1°	–	
		PDS	15–300	600	1.3°	–	
Chandra	99–?	ACIS-I	0.2–10	600	16'	0.5"	3200
		HRC	0.2–10	250	40'	0.5"	
		LETG/ACIS	0.2–10	100	–	0.5"	
		HETG/ACIS	0.4–10	40	–	0.5"	
XMM	99–?	EPIC-MOS	0.2–12	800	30'	6"	4800
		EPIC-PN	0.2–15	1000	30'	6"	
		RGS	0.3–2.5	150	–	6"	

Table 3.3 | (cont.)

Mission	Dates (launch- demise)	Detector	Energy range (keV)	Effective area (cm ²)	Field of view	Spatial resolution	Crab signal (cts s ⁻¹)
Swift	04-?	BAT	15–150	–	60° × 100°	17'	800
		XRT	0.2–10	110	23'	18"	
Suzaku	05-?	XIS	0.3–10	1360	18'	2'	2100
		HXD	15–200	200	34', 4°	–	

Chapter 17). It is designed to detect bursts, to quickly point instruments at the approximate position so an accurate location can be measured and to communicate this location to Earth. It observes the rapidly fading X-ray and optical afterglow while it is still relatively bright. It carries a burst alert telescope (BAT), an X-ray telescope (XRT), and a UV/optical telescope (UVOT). The coded-mask BAT views one-eighth of the sky and locates bursts with an accuracy of $\approx 4'$. Within a minute, sometimes faster, the other telescopes are pointed at this spot. The XRT gives a 5" burst location and the UVOT a 1" burst location relative to nearby stars. The XRT uses a CCD detector which gives the usual spectral information and can be used for research unrelated to γ -ray bursts. For Swift details, see Gehrels *et al.* (2004).

3.5.14 Suzaku: High-throughput spectroscopy

Suzaku (a mythological red bird) is a Japanese mission launched in July 2005. It carries four thin-foil telescopes focussed on CCD detectors, a fifth telescope focussed on a calorimeter and a mechanically collimated high-energy detector. It is in a low-Earth circular orbit with height 550 km.

There are five lightweight telescopes. Tightly nested thin-foil mirrors are in a Wolter type 1 configuration, but reflecting surfaces are approximated by cones. Reflection angles are small and efficiency is high for iron emission lines in the range 6–8 keV. Four telescopes focus on CCD arrays, the XISs, and the fifth on a calorimeter, the XRS. The XRS was a calorimeter, a revolutionary detector with spectral resolving power 10–25 times better than achievable with CCDs.

We all looked forward to seeing data from this instrument, which promised to change our way of thinking about X-ray spectroscopy of faint sources. Unfortunately, a Dewar venting problem caused the liquid He refrigerant to evaporate, and the calorimeter warmed before it could be used to record data from cosmic sources. The telescopes with CCDs are all operating nominally, and all have a particularly good response at low energies (≈ 0.5 keV).

A mechanically collimated hard X-ray detector, the HXD, covers the range 10–600 keV. This incorporates two types of detectors behind one collimator and in one anticoincidence shield and supplies broadband spectral coverage for the brighter sources. For Suzaku details, see Mitsuda *et al.* (2007).

References

Instruments, spacecraft and techniques

- Aschenbach, B. (1988) Design, construction, and performance of the ROSAT high-resolution X-ray mirror assembly, *Appl Opt* **27**, 1404.
- Bellazzini, R., *et al.* (2006) Gas pixel detectors for high-sensitivity X-ray polarimetry, *Proc SPIE* **6266**, 62662T.
- Boella, G., *et al.* (1997) BeppoSAX, the wide-band mission for X-ray astronomy, *A&A Supp* **122**, 299.
- Canizares C., *et al.* (2005) The Chandra high energy transmission grating: Design, fabrication, ground calibration and five years in flight, *PASP* **117**, 1144.
- CXC (2007), *Proposers' Observatory Guide*, TD 403.00.010, Cambridge, MA.
- Chodil, G., *et al.* (1967) Proportional counters and electronic systems for observing cosmic X-rays, *Rev Sci Inst* **38**, 1508.

- ESA (2009) http://xmm.esac.esa.int/external/xmm_user_support/documentation/uhb/node47.html.
- ESA/IAS/CEA (2009) http://integral.esa.int/integ_pictures.html#ibis.
- Evans, K. & K. A. Pounds (1968) The X-ray spectrum of a solar active region, *ApJ* **152**, 319.
- Forman, W. et al. (1978) The fourth Uhuru catalog of X-ray sources, *ApJ Supp* **38**, 357.
- Gehrels, N., et al. (2004) The Swift gamma-ray burst mission, *ApJ* **611**, 1005.
- Giacconi, R., et al. (1971) An X-ray scan of the galactic plane from Uhuru, *ApJ* **165**, L27.
- Giacconi, R., et al. (1979) The Einstein (HEAO-2) X-ray observatory, *ApJ* **230**, 540.
- Glasser, C. A., Odell C. E. & S. E. Seufert (1994) The proportional counter array (PCA) for the X-ray timing explorer satellite (XTE), *IEEE Trans Nucl Sci* **41**, 1343.
- Gruber, D. E. (1996) The high energy X-ray timing experiment on XTE, *A&A Supp* **120**, C641.
- Gursky, H., et al. (1966) A measurement of the location of the X-ray source SCO X-1, *ApJ* **146**, 310.
- Hill, J., et al. (2007) A burst chasing X-ray polarimeter, *Proc SPIE* **6686**, 66860Y.
- Jansen, F., et al. (2001) XMM-Newton observatory, *A&A* **365**, L1.
- Kelley, R. L., et al. (2007) The Suzaku high resolution X-ray spectrometer, *Publ Astron Soc Jpn* **59**, S77.
- Levine, A. M., et al. (1996) First results from the all-sky monitor on the Rossi X-ray timing explorer, *ApJ* **469**, L33.
- Meekins, J. F., et al. (1970) Solar soft X-ray flare spectra from OSO-4, *Sol Phys* **13**, 198.
- Mitsuda, K., et al. (2007) The X-ray observatory Suzaku, *Publ Astron Soc Jpn* **59**, S1.
- NASA/CXC (2010) http://chandra.harvard.edu/graphics/resources/illustrations/acis_nolabels.jpg.
- NASA/CXC/D. Berry (2010) <http://chandra.harvard.edu/graphics/resources/illustrations/cxcmirrors-72.jpg>.
- NASA/CXC/NGST (2010) http://chandra.harvard.edu/graphics/resources/illustrations/spacecraft_labeled-72l.jpg.
- NASA/MIT (2000) http://space.mit.edu/HETG/grat_cross-secs.gif.
- Palmieri, T., et al. (1975) Spatial distribution of X-rays in the Crab Nebula, *ApJ* **202**, 494.
- Pfeffermann, E., et al. (1987) The focal plane instrumentation of the ROSAT telescope, *Proc SPIE* **733**, 519.
- PIMMS (2008) <http://asc.harvard.edu/toolkit/pimms.jsp>.
- Rothschild, R. E., et al. (1974) Millisecond temporal structure in Cygnus X-1, *ApJ* **189**, L13.
- Seward, F. D. (1977) *Reflection Efficiencies of X-ray Mirrors 1–10 keV*, Report UCID-17505, Lawrence Livermore Laboratory, Livermore, CA.
- Serlemitsos, P. J., et al. The X-ray telescope with ASCA, *Publ Astron Soc Jpn*, **47**, 105.
- Sunyaev, R., et al. (1990) The ART-P and ART-S instruments, *Adv Space Res* **10**, 233.
- Tanaka, Y., H. Inoue & S. Holt (1994) The X-ray satellite ASCA, *PASJ* **46**, L37.
- Voges, W., et al. (1999) The ROSAT all-sky survey bright source catalog, *A&A* **349**, 389.
- Voges, W., et al. (2000) The ROSAT all-sky survey faint source catalog, <http://www.xray.mpe.mpg.de/rosat/survey/rass-fsc/>.
- Weisskopf, M. C., et al. (1978) A precision measurement of the X-ray polarization of the Crab Nebula without pulsar contamination, *ApJ* **220**, L117.
- Weisskopf, M. C., et al. (2002) An overview of the performance and scientific results from the Chandra X-ray observatory, *PASP* **114**, 1.
- White, N. & A. Peacock (1988) The EXOSAT observatory, *Soc Astron Ita Mem* **59**, 7.
- Wolter, H. (1952) Grazing-incidence mirror systems as image-forming optics for X-rays, *Ann Phys* **10**, 94.
- XMM_RGS (2008) http://xmm.esac.esa.int/external/xmm_user_support/documentation/technical/RGS.
- XMM_UH (2008) http://xmm.esac.esa.int/external/xmm_user_support/documentation/uhb_2.5/XMM_UH.

Reviews

- Bradt, H., T. Ohashi & K. Pounds (1992) X-ray astronomy missions, *Ann Revs Astron Appl* **30**, 391.
- HEASARC (2008) <http://heasarc.gsfc.nasa.gov/docs/observatories.html>.
- Seward, F. D. (2000) *Allen's Astrophysical Quantities*, 4th ed. edited by A. N Cox, 193, New York: AIP Press.

Chapter 4

Solar System X-rays

4.1 | The production of planetary X-rays

Planets are small and, compared to the cosmic subjects of other chapters, extremely weak sources of X-rays. Nevertheless, X-rays have now been detected from five planets, moons of Earth and Jupiter, several comets, and diffuse material in the solar neighborhood. These results have been scientifically useful and often surprising. The strongest X-ray source in the Solar System is, of course, the Sun. As in the visible band, orbiting solid objects shine with reflected solar energy. The soft X-ray luminosity of the solar corona is $\sim 4 \times 10^{27}$ erg s $^{-1}$, and that of the planets is a factor of $\sim 10^{14}$ weaker. Cometary X-rays are produced by collisions of energetic solar-wind particles with material in the comet. Some planets have magnetospheres which provide a mechanism for generating auroral X-rays. The energy that drives almost all these X-ray production processes originates in the Sun.

The observations are difficult, as targets move appreciably during the observation and are very bright optically; so bright that star sensors for aspect determination sometimes cannot be used. Soft X-ray detectors are also sensitive to visible light, which makes data reduction difficult. This chapter will cover Solar System objects in approximate order of X-ray detection.

4.2 | Earth

In some of the very first X-ray astronomy observations, solar X-rays scattered from the upper layers of the Earth's atmosphere were detected with rocket-borne proportional counters (Harries & Francey, 1968; Grader *et al.*, 1968). This X-ray glow from the sunlit Earth produced a strong, unpleasant background in the detectors, which had large fields of view. At energies below 1 keV, this background ruined the ability to detect faint cosmic sources. After this insight, rockets carrying soft X-ray detectors were launched at night, when the sunlit atmosphere was not visible to the detectors.

Most incident solar X-rays are absorbed through the photoelectric effect by atoms and molecules in the upper atmosphere. Only a small fraction are scattered, either by individual electrons or collectively by all the electrons bound in an atom (Thomson scattering). These scattered X-rays produce most of the signal in a detector looking down on the top of the sunlit atmosphere. The incident solar spectrum is quite soft, so the intensity of scattered X-rays is highest at low energies. There are also X-rays at specific energies from ionised atoms which recombine to emit fluorescent radiation.

The Earth's atmospheric composition up to 100 km is about the same as at sea level: 78 per cent

Table 4.1 Solar X-ray penetration of the atmosphere.

X-ray energy (keV)	1	5	20
Minimum height (km)	100	75	50

molecular nitrogen, 21 per cent molecular oxygen and 1 per cent argon. Above 125 km atomic oxygen is the main constituent; above 1000 km, helium; above 2500 km, atomic hydrogen. Solar X-rays penetrate to an altitude that is strongly dependent on the X-ray energy (see Table 4.1). The layer of the atmosphere seen by a detector above it depends on the energy band of the detector. In the 0.2–1 keV band, observable scattering occurs at altitudes of 100–120 km; and in the 1–10 keV band, at altitudes of 70–100 km.

4.2.1 Auroral X-rays

Auroral particles are a second source of atmospheric X-rays. Electrons precipitating into the atmosphere (at locations determined by the Earth's magnetic field) generate bremsstrahlung and some fluorescence. This precipitation occurs in the auroral zones. On Earth these form two ovals at 65° to 70° magnetic latitude, circling the north and south magnetic poles. Here the magnetic field lines connect with the boundaries of the radiation belts where energetic electrons are stored. The precipitating electrons are not very energetic, and the optical aurora occurs at altitudes of 100–300 km. None of the electrons penetrate to altitudes less than 80 km.

X-rays generated by these particles, the auroral-electron bremsstrahlung, have a soft spectrum. If viewed from below, most X-rays are absorbed by the atmosphere. There is, however, a hard X-ray component with energy above 25 keV which can penetrate to 40 km, and this altitude can be reached by balloons. For several years starting in 1957, balloon-borne detectors were used to measure auroral X-rays (e.g. Anderson & Milton, 1964). The flux, spectrum and time dependence of precipitating electrons were deduced from these observations.

Auroral X-rays should also be detectable from a point directly above the source. The measurement was first attempted using detectors carried

by sounding rockets or near-Earth satellites and was unusually difficult. In these circumstances, at the time of making such a measurement, the instrument would be embedded in the flux of electrons that are causing the aurora. These generate X-rays within the detector which overwhelm the atmospheric X-rays of interest. Modern spacecraft, however, positioned well beyond the magnetosphere and viewing the Earth from a distance, can easily see auroral X-rays and have now measured the time and spatial dependence of emission. By comparing these data with the velocity and density of the solar wind which pushes on the magnetosphere, the auroral mechanism can be studied. Colour Plate 7 shows a map of auroral X-rays from the PIXIE instrument on the Polar spacecraft. Colour Plate 8 shows a high-resolution Chandra image.

4.2.2 Scattering and fluorescence

The intensity of solar X-rays scattered from the atmosphere is quite weak, except during flares, when solar X-ray emission is strong. The following example is typical and shows the effect of viewing geometry. The British satellite Ariel V carried a proportional counter with a 3.5° field of view, sensitive in the energy range 1.5–10 keV and pointed along the satellite spin axis. The detector was pointed at various astronomical targets by moving the spin vector and holding it steady for the duration of each observation. Since the spacecraft was in a near-Earth orbit, the Earth usually passed through the field of view once every orbit. Usually the Earth was dark, and no X-rays were detected from the atmosphere day or night.

On 28 March and 30 April 1976, there were two large solar flares. During these flares the X-ray flux from the Sun greatly increased. The Earth, as seen by the Ariel detector, also brightened. At the start of the flare the Earth-scattered flux increased by a factor of 1000 and then decayed over the next 6 hours in a way which exactly followed the solar X-ray emission observed by another satellite. During each orbit the Earth was viewed for 40 min. Within this interval the X-ray flux varied much more rapidly than did solar emission because the illuminated Earth is limb-brightened, an effect of absorption and viewing conditions.

Box 4.1 Scattering geometry

A simple calculation, illustrated in Fig. 4.1, shows this dependence. Assume that the top of the atmosphere is planar, the X-ray absorption coefficient is the same for both incident and scattered X-rays, there is only a single scattering and solar flux is constant during the observation. If the angle between the path of the incident X-ray and the terrestrial radius vector is θ and the angle between the path of the scattered X-ray and the radius vector is ϕ , then the scattered flux at the satellite, F_s , is

$$F_s \propto [1 + (\cos \phi / \cos \theta)]^{-1}.$$

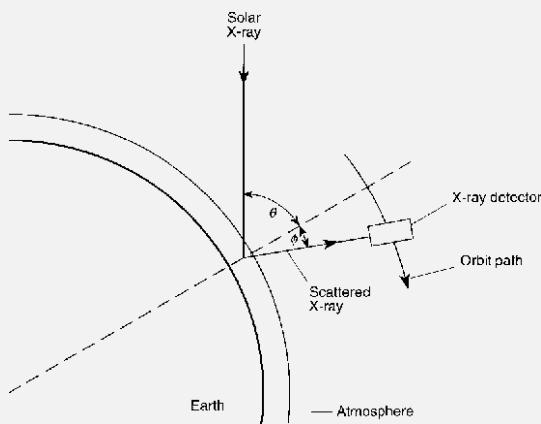


Fig 4.1 Sun-Earth-satellite configuration during the Ariel V observations of scattered solar X-rays. The observed scattered flux is shown in Fig. 4.2

This function was used to calculate the solid line in Fig. 4.2, and it fit the measurements well. The detector window was too thick to transmit X-rays from nitrogen or oxygen, but the argon fluorescent line at 3.0 keV was prominent in the spectrum.

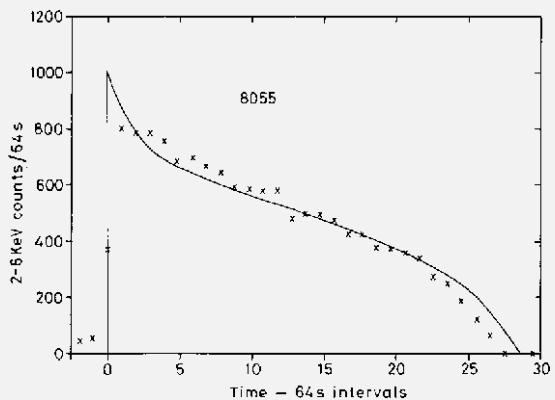


Fig 4.2 Counting rate of the Ariel V X-ray detector during orbit 8055. The sunlit Earth entered the field of view at time 0, and by time 30 the detector was viewing the dark Earth. The solid line shows the predicted dependence of scattered X-rays on satellite position. The decrease in scattered flux is from increased absorption in the atmosphere as the geometry changes. When the Earth is viewed perpendicular to the direction of illumination, there is considerable limb brightening (from Seward *et al.*, 1976).

The scattered flux at the satellite depended on the orientation of the Sun, Earth and detector. Because the satellite spin-axis remained fixed in space, the field of view scanned over the top of the atmosphere as the satellite circled the Earth. The angle that detected scattered X-rays made with the terrestrial radius vector changed. The total thickness of atmosphere that these X-rays traversed, traveling from Sun to Earth to detector, also changed. The thicker this layer, the more the X-rays were attenuated. Thus the detected flux depended strongly on the position of the satellite in its orbit.

The maximum observed flux occurred when the detector viewed the sunlit limb of the Earth and was looking parallel to the top of the atmosphere. When the scattered flux dropped rapidly to zero, the detector was looking at the terminator (the dividing line between day and night). At this point solar X-rays were absorbed before they reached that part of the atmosphere viewed by the detector. This same limb brightening was seen with the Einstein telescope, also in a near-Earth orbit (Fink *et al.*, 1988). The Einstein instruments, however, were sensitive to softer X-rays than Ariel V, and scattered X-rays from the quiet

Box 4.2 | Scattered solar X-rays

Armed with the albedo measured from Earth's atmosphere during a solar-quiet period, we can roughly calculate the expected signal from scattered/fluorescent X-rays from other planets. Table 4.2 gives count rates predicted for the Einstein IPC (≈ 0.5 ROSAT PSPC, ≈ 0.2 Chandra ACIS). Assumptions were that all objects have an albedo of 10^{-4} and that the phases of Mercury, Venus and

the Moon were 0.5. Predicted X-ray luminosities are given and compared with measured values. To illustrate the amusingly weak nature of these sources, units are megawatts (MW), the same as for Earth-based power plants. Data are from Bhardwaj *et al.* (2007). Solar activity and/or particle bombardment can cause the measured L_x to be an order of magnitude more than these predictions.

Table 4.2 | Predicted counting rates and luminosities for planets.

Object	Solar flux relative to Earth	Solid angle at Earth (arcmin 2)	Predicted rate (counts s $^{-1}$)	Predicted L_x (MW)	Observed L_x (MW)
Mercury	6.6	0.005	0.0016	18	-
Venus	1.9	0.03	0.003	32	50
Earth	1.0	1	0.05	18	40
Moon	1.0	360	18	1	0.07
Mars	0.43	0.07	0.0013	2	2
Jupiter	0.037	0.48	0.001	90	1000
Io	0.037	0.0003	5×10^{-7}	0.05	2
Saturn	0.011	0.08	0.00004	18	200

Sun were sometimes a considerable background problem.

4.3 | The Moon

Having been the principal target at the time of the first successful detection of a cosmic X-ray source, it is interesting to note that X-rays from the lunar surface were only first observed in June 1990 with ROSAT. Ten years earlier, Einstein could easily have detected the X-rays but did not try because of potential damage to its star-tracking camera. This is not a trivial observation. The Moon is not only optically bright but a rapidly moving target. Either it must be tracked during the observation or the image must be reconstructed to follow the motion after the fact. The apparent motion with respect to the stars is both the lunar orbital motion of ≈ 30 arcmin hour $^{-1}$ and, for an instrument in near-Earth orbit, an even

larger parallax. In theory an X-ray picture of the full Moon might show bright and dark spots because of differences in surface composition, but such differences in X-ray albedo are expected to be small.

Figure 4.3 shows the ROSAT picture. Schmitt *et al.* (1991) assembled this from a 1500-s interval during which the Moon passed from the edge to the centre of the field of view. The Sun was illuminating the Moon from the right side of the picture, and X-rays scattered from the illuminated side were brightest at the limb, as expected. Total luminosity of these X-rays was 7×10^{11} ergs s $^{-1}$, the power needed to operate 700 100-W light bulbs! In another historic first, the diffuse X-ray background was observed to be shadowed, showing that it originates (as all believed) beyond the lunar orbit. However, the count rate observed from the lunar dark side was significantly higher than the detector's internal background. These extra events are thought to be charge-exchange

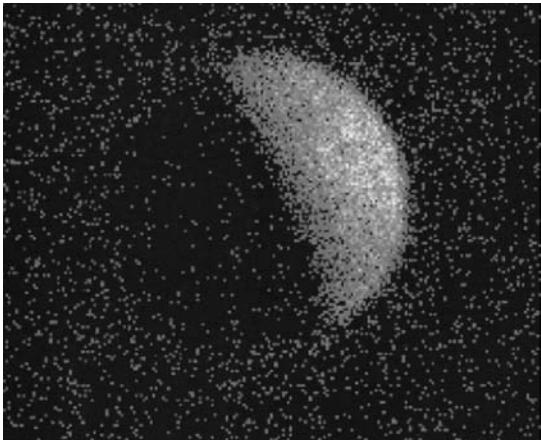


Fig 4.3 The Moon observed by ROSAT. X-rays scattered from the solar-illuminated surface are brightest at the limb and fade to nothing at the terminator. The Sun-Moon-detector orientation was as illustrated in Fig. 4.1. Note that the Moon shadows the diffuse X-ray background so the dark side of the Moon shows less emission than the surrounding sky (from Schmitt *et al.*, 1991).

X-rays generated by collisions of solar wind ions with geocoronal H atoms in the space between Earth and the Moon.

4.3.1 What is the lunar surface made of?

The ROSAT picture of the Moon's bright side was a first step towards using X-ray imaging from a distance for lunar surface diagnostics. With this purpose, the Moon has been observed several times by Chandra. Fluorescent emission from O, Mg, Al and Si was detected (Wargelin *et al.*, 2004), indicating that fluorescence is the dominant production mechanism. There was evidence for differences in albedo between highlands and maria. Small differences are expected because of different elemental compositions. The goal of such observations is a map of lunar crust elemental composition. To obtain this, long observations will be needed, even during the time of solar maximum, when quiescent solar flux is highest.

Lunar crust diagnostics were actually accomplished in 1971 and 1972 by taking the detectors much closer to the Moon (Adler *et al.*, 1973). A special set of X-ray detectors carried in lunar orbit by the Apollo 15 and 16 command modules was used. A map of scattered X-rays from the sunlit lunar

Box 4.3 | The dark side of the Moon

The dark side of the Moon does not appear completely dark to an X-ray detector because of weak emission along the path from Earth to Moon. Earth's atmosphere at altitudes above 2500 km is mostly hydrogen, and the density falls as approximately the inverse cube of the distance from Earth. At 10 Earth radii distance the density is $\sim 25 \text{ atoms cm}^{-3}$. High-velocity ions in the solar wind collide with this H and can capture the electrons. Thus an O^{+8} ion in the wind can capture an electron becoming O^{+7} in an excited state. Subsequent decay to the ground state produces X-rays characteristic of hydrogen-like O. The cross section for this process is large and accounts for X-ray production in the geocorona and from comets. The solar wind velocity varies between 300 and 500 km s^{-1} . When it increases, Earth's magnetosphere is compressed, and solar-wind ions can approach closer to Earth. Because neutral H is not affected by the magnetosphere and the H density goes as r^{-3} , there are more target atoms, and X-ray production increases greatly. This variability has been used to identify solar-wind charge exchange with geocoronal H as the source of these background X-rays.

surface was assembled from approximately 200 hours of data. These were taken at an altitude (above the Moon's surface) of 100 km. Although restricted to sunlit territory directly under the spacecraft, 20 to 25 per cent of the lunar surface was examined during the two missions.

The instrumentation consisted of three proportional counters operating in the energy range 0.7–2.8 keV. Each was pointed at the surface and collimated to a 30° field of view. There was also a solar monitor. Two of the counters were covered with thin foils of Mg and Al, respectively. The K absorption edges of these materials formed transmission windows which enhanced the relative response at particular energies. One detector had relatively high efficiency for the fluorescent $K\alpha$ line from Mg. The other had a relatively high efficiency for the $K\alpha$ line from Al. The third detector was equally sensitive to fluorescent lines from Mg, Al and Si. The relative

counting rates of these three detectors were used to determine the relative abundances of the three elements in the lunar material. Because the solar X-rays are easily absorbed, the measurement only applied to a depth of 0.1 mm, truly the lunar surface!

The spatial resolution was 100 km. It was inferred that there exists a global lunar crust characterised by a high Al-to-Si ratio. This occurs in the highland areas both on the near and far sides. The highland areas show high Al and low Mg concentrations, whereas the reverse is true in the maria, with the irregular maria showing a higher Al-to-Si ratio than the circular maria. Comparison with optical pictures showed that areas of high optical albedo correspond to areas of high Al-to-Si composition. Because two of the Apollo landing sites were surveyed, lunar samples returned to Earth were used to calibrate the system. The X-ray data were thus tied to actual samples of lunar material. Other missions to the Moon will attempt more detailed mapping with similar instruments. This technique has also been used to study the composition of the asteroid Eros with the NEAR-Shoemaker spacecraft (Nittler *et al.*, 2001)

4.4 Comets

Most comets have orbits with large eccentricity and spend almost all their time at large (several astronomical units) distances from the Sun. When cold, a typical comet is a solid chunk of debris from the early Solar System – ice, dust and rock, with diameter of a few kilometers. When close to the Sun, this material is warmed, and its gas and dust are evaporated, causing the solid nucleus to become surrounded by a cloud of gas with a diameter up to 10^6 km. This extended volume is referred to as the comet's coma. The solar wind streams past the comet and pushes gas and dust away from the coma, forming the two tails which are the hallmark of comets. The ion tail is carried away from the Sun by the solar wind and can be visible for a length of ≈ 1 AU (10^8 km). The dust tail is not as long and follows the track of the comet. During the Einstein mission, Hugh Hudson proposed several times to search for X-rays from a comet.

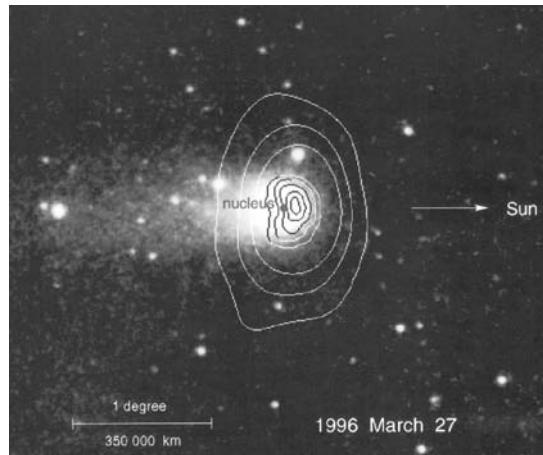


Fig 4.4 ROSAT image of Comet C/1996 B2 Hyakutake. Contours of constant X-ray brightness are superposed on an optical picture of the comet. X-rays are produced by solar-wind ions colliding with gas in the coma, which surrounds the tiny nucleus (MPE/ROSAT 1997; Lisse *et al.*, 1996).

Since it was generally believed that cometary X-ray emission was close to zero, the first proposals were not accepted. Late in the mission, time was grudgingly given for a short observation of Comet Bradfield (1979). Inspection of the 2.5-ks observation revealed no obvious source, and an upper limit of 10^{14} erg s $^{-1}$ was obtained (Hudson *et al.*, 1981). No other observations were attempted until 1996 when, in the middle of the ROSAT mission, Lisse *et al.* (1996) discovered X-rays from Comet C/1996 B2 Hyakutake. Figure 4.4 shows this result.

X-ray emission was strong, and with this inspiration, the 1990–1991 ROSAT all-sky survey was searched for emission from other comets (Dennerl *et al.*, 1997). Four comets were found to have been detected on seven occasions. Thus an astronomical surprise opened a field of study, and at present, dozens of comets have been observed by several X-ray observatories. The X-ray luminosity varies from 10^{14} to 10^{16} erg s $^{-1}$. X-rays come from the coma, so the emitting region is diffuse. The maximum is a crescent-shaped region on the sun side of the nucleus, and emission fades into the background an arcminute or two away. Observing strategies for comets are complicated. Although X-ray emission is greatest

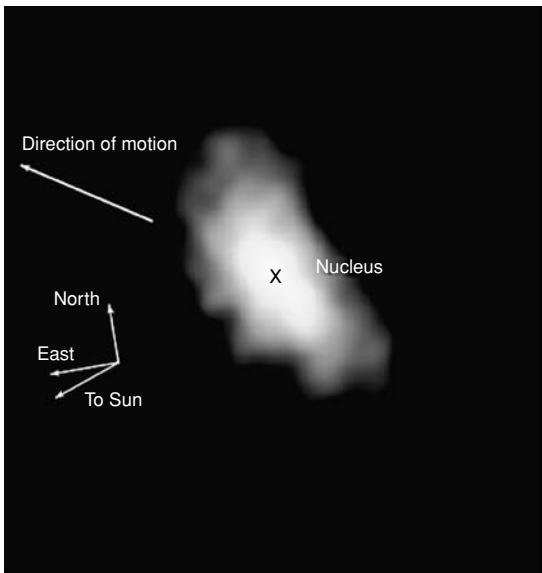


Fig 4.5 Chandra ACIS image of Comet C/1999 S4 (LINEAR). The field shown is 10' square and the exposure was 3 hours. X-ray emission is diffuse and extends 4' from the bright area close to the nucleus. Emission is symmetrical about the comet–Sun line (NASA/CXC 2000; Lisse *et al.*, 2001).

when the comet is closest to the sun, the brightness also depends on distance from the spacecraft. Observations are accomplished in a series of short observations (because the target is rapidly moving) around a time which is a balance of the minimum solar distance and the minimum spacecraft distance. Figure 4.5 shows the Chandra observation of Comet C/1999 S4 (LINEAR), which was one of the brighter comets and which was discovered by the Lincoln Near-Earth Asteroid Research project.

Comet observations are all ‘photon starved’, as they are weak sources and observation times are limited. Nevertheless, good spectra have been obtained which show that the production mechanism is solar wind charge exchange. Figure 4.6 is a comet spectrum from the ACIS instrument on Chandra which shows that the energy is predominantly in emission lines of several solar-wind ion species. These ions collide with cometary neutral atoms in the coma and pick up an electron, and the resulting cascade to the ground state produces X-rays with energies characteristic of the ion. Because fluorescent solar X-rays would have

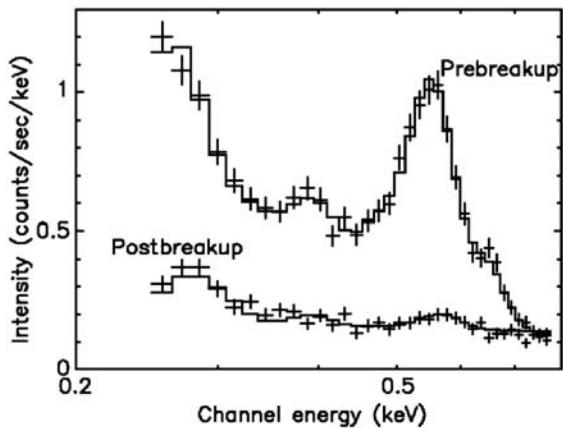


Fig 4.6 Chandra ACIS spectrum of Comet C/1999 S4 (LINEAR). Solid line shows a fit with six lines produced by charge exchange between material from the comet and energetic ions in the solar wind. The principal peak is from 0.56 keV photons from O⁺⁷. After the first observation, the comet broke up, and X-ray emission was a factor of 4 less during the second observation (from Lisse *et al.*, 2001).

slightly different energies, characteristic of neutral atoms, we know that the contribution of fluorescent X-rays is small, as is continuum radiation expected from electron bremsstrahlung. So, because X-ray production depends almost solely on properties of the solar wind, comets have been proposed as useful tools to study the solar wind at otherwise inaccessible locations.

4.5 Jupiter

In 1955 Jupiter was identified as a strong source of nonthermal radio emission (Burke & Franklin, 1955). This was soon recognised to be synchrotron radiation, establishing the presence of a magnetosphere with a large population of energetic electrons. By analogy with Earth, these electrons were expected to precipitate along field lines and to generate X-rays when they collided with atoms in the Jovian atmosphere. There was also speculation that bombardment of the inner Galilean satellites might produce characteristic X-rays from their surfaces.

In 1974 the spacecraft Pioneer 10 and 11 measured the extent of the radiation belt surrounding

Jupiter (Van Allen *et al.*, 1975). The zone of stable trapping extends to 25 times the radius of the planet. This is big! The zone encompasses the orbits of Amalthea (a small moon), Io, Europa and Ganymede. The trapped particle flux was found to decrease at the orbital distance of Amalthea and Io. Io, in fact, almost completely removes low energy protons from the radiation belts.

Auroral radiation from Jupiter was detected in the UV range by instruments on the Voyager spacecraft. Subsequent measurements with the IUE spectrometer showed the radiation to consist primarily of hydrogen Lyman α at 1216 Å and the 1175–1650 Å Lyman and Werner bands of molecular hydrogen. Hydrogen is, of course, the major constituent of the upper atmosphere of Jupiter. X-rays from Jupiter were first detected on 13 April 1979 by Metzger *et al.* (1983) using the Einstein IPC. Jupiter was observed four times, each observation chosen to be when retrograde motion was starting or ending. In this way the apparent motion of the planet during the observation (typically of 3 hours' duration) was kept to a minimum of $\approx 1'$. The first observation was quickly inspected, and no emission was seen at the centre of the field. Before cancelling the remaining observations, the first was re-examined ‘one last time’, and the pointing direction was found to be incorrect! The 1979 coordinates of Jupiter had not been precessed to 1950, which was the epoch required for Einstein target coordinates. There was indeed a source 20' off-axis at the actual position of Jupiter. Subsequent IPC observations confirmed this detection and showed the flux to be relatively constant and to have a very soft spectrum. The Einstein HRI showed that the emission was from the polar regions and consequently auroral in nature. This satisfying result ended a 15-year search which started with the first observations of extrasolar X-rays.

4.5.1 High-resolution imaging of Jupiter

High-resolution observations with Chandra and XMM-Newton have since established the location and spectrum of Jupiter's auroral X-rays. Because the planet is a moving, spinning target, and a rather weak source, this is a non-trivial accomplishment. The X-ray aurora was found to occur

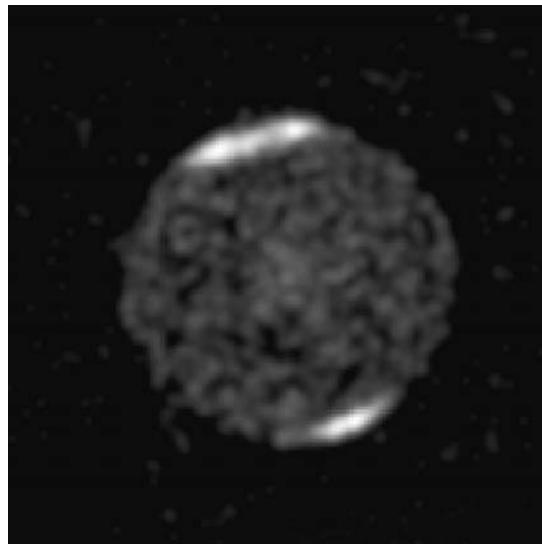


Fig 4.7 A Chandra HRC image of Jupiter. There is bright auroral emission from both poles and relatively uniform emission from the disc. The diameter of the disc is 45'' (from Elsner *et al.*, 2005).

at a higher latitude than the UV aurora, and magnetic field lines at the latitude of the X-ray emission connect to Jupiter's outer magnetosphere. The spectrum has an energetic component indicating electron bremsstrahlung and line emission from O and perhaps C (Branduardi-Raymont *et al.*, 2006). These heavy ions are either from the solar wind or are precipitated from the magnetosphere, perhaps both. Rapid time variability, including one detection of a pronounced 40-min period, strengthens this conclusion, although the periodicity is not understood.

There is also emission from Jupiter's disc. This is relatively uniform (Fig. 4.7), and the spectrum is as expected from the scattering of solar X-rays. A variation with time which tracks solar emission proves that this is indeed the origin of most disc X-rays. A contribution from magnetospheric heavy ions, however, is not ruled out.

The Jupiter system is a rich (meaning many processes, most not well understood) source of emission. Jupiter's moons are separated from the planet by several arcminutes and so are easily resolved. Emission from Io and Europa has indeed been detected by Elsner *et al.* (2002). Approximately 10 photons were collected from each moon

in an 86 ks Chandra ACIS observation, a luminosity considerably higher than expected from solar X-ray fluorescence. The moons are immersed in the radiation belts, and spectra (very limited) are consistent with heavy ion bombardment. Yet another source is a plasma torus associated with Io's orbit. The torus is defined by optical emission from O and S ions and has been mapped with terrestrial telescopes. The origin of these elements is thought to be the volcanic plumes of Io. Atoms are first spewed into space by the volcanos then ionised and trapped by Jupiter's magnetic field. This torus is a faint diffuse source of soft X-rays. Lastly, Jupiter was observed during the impact of Comet Shoemaker-Levy in 1994 (Waite *et al.*, 1995). Auroral X-rays were associated with the impact of the fragments. Although the electromagnetic forces which accelerate particles in the magnetosphere are not well understood, the rotation of Jupiter is probably a significant source of power for the high-energy ions and electrons which produce these Jovian X-rays.

4.6 | Saturn

X-rays from Saturn were first observed in 2002 by XMM-Newton (Ness *et al.*, 2004) and verified in subsequent observations. Emission is from the disc, and in contrast to Jupiter, auroral emission is not strong. Emission is variable and has been observed to increase at the time of a solar flare. So scattering and/or fluorescence of solar X-rays is the principal production mechanism. There is also weak emission from the rings, probably from particle bombardment.

Saturn has a large moon, Titan, whose diameter is 0.40 that of Earth, and it has a thick atmosphere consisting of nitrogen, argon and methane. On 5 January 2003 Titan passed in front of the Crab Nebula, a once-in-a-lifetime event. During the transit, Chandra observed for ≈ 30 ks. The image was carefully reconstructed to track the 1" diameter moon as it traversed the $\approx 90''$ bright X-ray centre of the Crab (Fig. 9.4). Figure 4.8 shows the resultant image of Titan's shadow. The measured thickness of Titan's atmosphere was 880 km, 17 per cent of its radius (Mori *et al.*, 2004) – an observational tour de force.

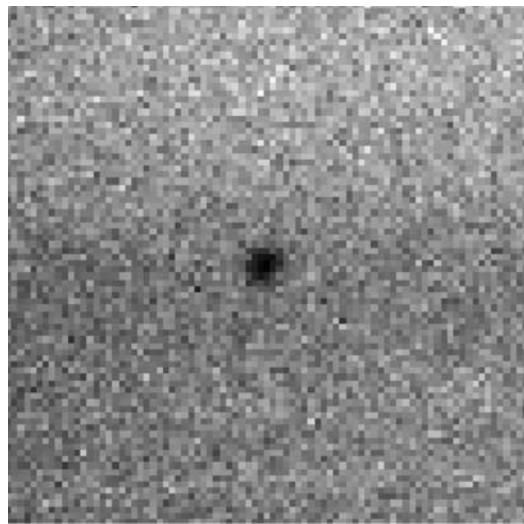


Fig 4.8 The X-ray shadow cast by Titan passing in front of the Crab Nebula as observed by Chandra. The image is 20" in extent, and each pixel is 0.25". Details in the bright Crab are smeared by the tracking of the moving moon. The shadow is not sharp edged because of the absorption in Titan's atmosphere (from Mori *et al.*, 2004).

4.7 | Venus and Mars

The atmospheres of both Venus and Mars consist predominantly of carbon dioxide. Although the atmospheric pressure on Venus is 100 times that of Earth and the pressure on Mars is only 1/100th, both atmospheres are thick to X-rays, and the scattering/fluorescence of solar X-rays is observed only from their upper layers. X-rays from Mars were first detected in July 2001 by Dennerl (2002) using the Chandra ACIS-I detector. The signal was only $\sim 10^{-2}$ counts s^{-1} , with emission being observed from the planetary disc and from a faint halo (Fig. 4.9). A second observation in November 2003 was accomplished by XMM-Newton (Dennerl *et al.*, 2006). The diffraction gratings (RGS) produced a spatially resolved high-dispersion spectrum. Although the source is weak, almost all emission is in discrete spectral lines which stand out from the background. Emission from Mars was found to have two components: fluorescent radiation from the planetary disc and emission from solar-wind charge exchange during collisions with atoms in the exosphere or

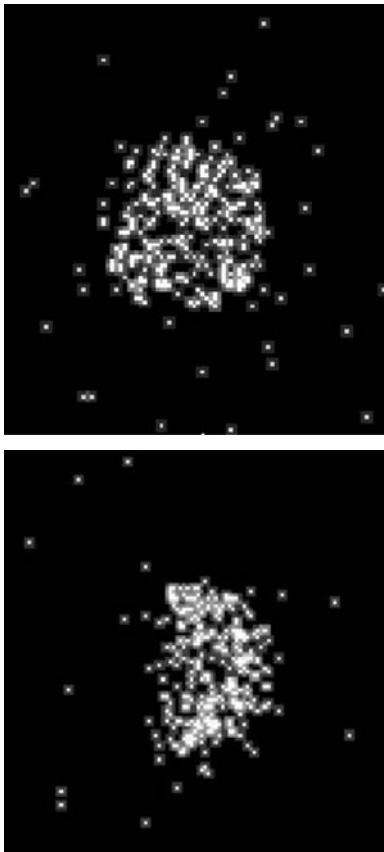


Fig 4.9 The discovery of X-rays from (top) Mars and (bottom) Venus. Solar X-rays cause fluorescence in the upper atmosphere, and we see the sunlit part of the discs. Each bright dot shows the origin of one X-ray photon in the energy range 0.4–0.7 keV. There is no smoothing. These are not bright sources. The Martian halo is too faint and large to be apparent in this figure. The planetary discs subtended angles of 20'' and 23'', respectively. There is some limb brightening in the Venus image because of viewing geometry (from Dennerl 2002, Dennerl *et al.*, 2002).

halo. The exosphere was found to extend up to 8 Mars radii (50'' from the planet). X-ray luminosity of the disc was 4 MW, 3 times that of the first observation in 2001, as a result of increased solar activity. X-ray luminosity of the halo was 13 MW.

Figure 4.10 shows the measured spectrum of the disc. The two dominant lines are from neutral O atoms in CO₂. The O K α line is split by an additional energy state in the molecule, and there is a weak line from N₂. All other weak lines are from

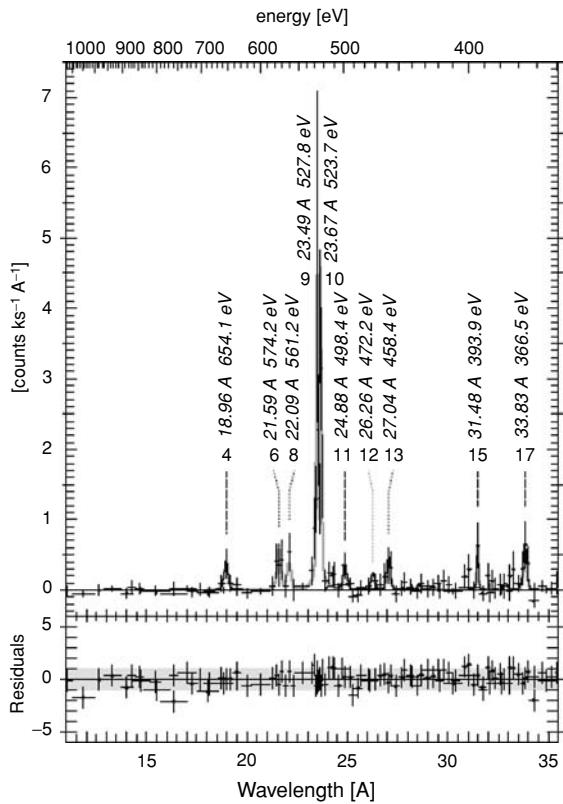


Fig 4.10 XMM-Newton RGS spectrum of the Martian disc. The two strongest lines are from fluorescence by O in the CO₂ molecule. Solid lines are a model fit to the data (from Dennerl *et al.*, 2006).

highly charged solar wind ions picking up an electron from material in the exosphere close to the planet, mostly C⁺⁴, C⁺⁵, O⁺⁶ and O⁺⁷. The relative line intensities from He-like O⁺⁶ show that this ion is not produced through collisional equilibrium in the atmosphere.

Venus was first observed by Chandra in 2001. Figure 4.9 shows the result. The X-ray image is of the solar-illuminated disc, similar to Mars, but Venus is limb brightened. The dominant emission is solar X-ray fluorescence. Indeed, data taken with the Chandra ACIS/LETG high-resolution spectrograph show only C, N and O K α lines. No auroral brightening was observed, and any halo from solar-wind charge exchange in the exosphere was expected to be too faint to show in this observation.

Thus soft X-rays have been observed from (astronomically speaking) small, cold bodies in

our solar system. The energy contained in solar X-rays and in the solar wind drives the emission of most objects. The Earth and Jupiter, however, have strong magnetic fields, and magnetospheric processes account for relatively bright and variable emission. Observations of planets required very careful planning, whereas several comets were serendipitously present in a survey – something for everyone.

Box 4.4 Observing moving targets from space

The Dennerl *et al.* (2006) paper describes special procedures required for observation of a moving target. The viewing of XMM-Newton is restricted to lie in a band $\pm 20^\circ$ from the spacecraft–Sun line, an elongation between 70° and 110° , so Mars could not be observed at opposition when it was closest to Earth. The 2003 observation lasted 10^5 ks, during which time Mars moved $35'$ with respect to the background sources. The time was divided into 12 short pointings to keep the planet close to the centre of the field.

The logistics of the Venus observation are also interesting. The maximum separation between Venus and the Sun does not exceed 47.8° . This was too close to the Sun to be feasible for observation by Einstein, ROSAT or XMM-Newton, all restrained to observe approximately perpendicular to the Earth–Sun line. Chandra, however, could observe to within 45° of the Sun, thereby opening a narrow window of opportunity. Observations were possible only during a 5-week interval every ~ 3 years. With careful scheduling, Venus was observed in January 2001 and March 2006.

References

Reviews

Bhardwaj, A., *et al.* (2007) X-rays from Solar System bodies, *Planet Space Sci* **55**, 1135.

Earth and Moon

Adler, I., *et al.* (1973) Apollo 15 and 16 results of the Integrated Geochemical Experiment, *Moon* **7**, 487.

- Anderson, K. A. & D. W. Milton (1964) Balloon observations of X rays in the auroral zone, 3, High time resolution studies, *J Geophys Res* **69**, 4457.
- Bhardwaj, A., *et al.* (2007b) First terrestrial soft X-ray auroral observation by the Chandra X-ray Observatory, *ASTP* **69**, 179.
- Fink, H., J. Schmitt & F. R. Harnden Jr. (1988) The scattered X-ray background in low-Earth orbit, *Astron Astrophys* **193**, 345.
- Grader, R. J., R. A. Hill & F. D. Seward (1968) X-ray airglow in the daytime sky, *J Geophys Res* **73**, 7149.
- Harries, J. R. & R. J. Francey (1968) Observation of Cen XR-2, Sco XR-1, and Terrestrial X-rays, *Aust J Phys* **21**, 715.
- NASA-GGS/LMMSC (2008) <http://pixie.spacsci.com/>.
- NASA/MSFC/CXC (2005) <http://chandra.harvard.edu/photo/2005/earth/>.
- Schmitt, J. H. M. M., *et al.* (1991) A soft X-ray image of the Moon, *Nature* **349**, 583.
- Seward, F. D., *et al.* (1976) Argon fluorescent X-rays in the Earth's atmosphere during solar flares, *Nature* **264**, 421.
- Wargelin, B. J., *et al.* (2004) Chandra observations of the 'dark' Moon and solar wind charge transfer, *ApJ* **607**, 596.
- Venus and Mars**
- Dennerl, K. (2002) Discovery of X-rays from Mars with Chandra, *A&A* **394**, 1119.
- Dennerl, K., *et al.* (2002) Discovery of X-rays from Venus with Chandra, *A&A* **386**, 319.
- Dennerl, K., *et al.* (2006) First observation of Mars with XMM-Newton: High resolution X-ray spectroscopy with RGS, *A&A* **451**, 709.
- Jupiter and Saturn**
- Burke, B. F. & K. L. Franklin (1995) Observations of a variable radio source associated with the planet Jupiter, *J Geophys Res* **60**, 213.
- Branduardi-Raymont, G., *et al.* (2006) XMM-Newton observations of X-ray emission from Jupiter in *Proceeding of Symposium: The X-ray Universe 2005*, 15, Madrid: El Escorial.
- Elsner, R. F., *et al.* (2002) Discovery of soft X-ray emission from Io, Europa, and the Io plasma torus, *ApJ* **572**, 1077.
- Elsner, R., *et al.* (2005) Simultaneous Chandra X-ray, Hubble Space Telescope ultraviolet, Ulysses radio observations of Jupiter's aurora, *J Geophys Res* **110**, A01207.
- Metzger, A. E., *et al.* (1983) The detection of X-rays from Jupiter, *J Geophys Res* **88**, 7731.

- Mori, K., et al. (2004) An X-ray measurement of Titan's atmospheric extent from its transit of the Crab Nebula, *ApJ* **607**, 1065.
- Ness, J. U., et al. (2004) X-ray emission from Saturn, *A&A* **418**, 337.
- Van Allen, J. A., et al. (1975) Pioneer 11 observations of energetic particles in the Jovian magnetosphere, *Science* **188**, 459.
- Waite, J. H., Jr., et al. (1995) ROSAT observations of X-ray emissions from Jupiter during the impact of comet Shoemaker-Levy 9, *Science* **268**, 1598.
- Hudson, H. S., W.-H. Ip & D. A. Mendis (1981) An Einstein search for X-ray emission from comet Bradfield 1979 1, *Planet Space Sci* **29**, 1373.
- Lisse, C., et al. (1996) Discovery of X-ray and extreme ultraviolet emission from comet C/Hyakutake 1996 B2, *Science* **274**, 205.
- Lisse, C., et al. (2001) Charge exchange-induced X-ray emission from comet C/1999 S4 (LINEAR), *Science* **292**, 1343.
- MPE/ROSAT (1997) <http://www.mpe.mpg.de/xray/wave/rosat/gallery/calendar/1997/mar.php>.
- NASA/CXC (2000) <http://chandra.harvard.edu/photo/2000/c1999s4/>.
- Nittler, L. R., et al. (2001) X-ray fluorescence measurements of the surface elemental composition of asteroid 433 Eros, *Meteor Planet Sci* **36**, 1673.

Comets and asteroids

Dennerl, K., J. Englhauser & J. Trumper (1997) X-ray emissions from comets detected in the Rontgen X-ray satellite all-sky survey, *Science* **277**, 1625.

Chapter 5

X-ray absorption and scattering in the Interstellar Medium

5.1 | The Interstellar Medium

The space between stars is not empty. It is full of gas and dust which are collectively called the *Interstellar Medium* or ISM. The ISM accounts for ~ 10 per cent of the mass of our Galaxy. To see anything beyond the Solar System, we must look through the ISM, and thus all observations are filtered and modified. In our Galaxy the gas forms a disc in the plane of the Milky Way, with diameter ≈ 30 kpc and thickness ≈ 0.7 kpc. The density averages about 1 atom/cm³, a far better vacuum than any that could be created here on Earth. This does not sound like much, but it is enough to absorb soft X-rays from most galactic sources. The composition of the gas is close to the usual cosmic abundance: 90 per cent H (by number), 10 per cent He and 0.1 per cent heavier elements. However, it is far from being a uniform medium. The neutral gas exists in a very large range of density, n , and temperature, T . A diffuse cloud might have $n \sim 100$ and $T \sim 80$ K. The medium between clouds might have $n \sim 1$ atoms cm⁻³ and $T \sim 8000$ K. There is also warm (8000 K) and hot (300 000 K) ionised material. Our interest here is in the neutral gas, which does most of the absorbing.

Neutral H in the ISM has, for more than 60 years, been directly observed in the radio band at a wavelength of 21.106 cm. In the lowest state of a H atom the spin of the proton and electron are either parallel or antiparallel. A change from the parallel to antiparallel state is called a *spin flip transition*, in which a quantum of radiation with this very

Box 5.1 | Interstellar dust

The dust in the ISM is thought to consist of small carbonaceous and silicate grains. Because light is scattered by small particles, this dust causes extinction of the light from stars. Stars at a distance of 1 kpc are about 1 magnitude dimmer than expected because of this extinction. Because blue light is scattered more strongly than red light, there is also a reddening of the star's spectrum. Measured UV/optical extinction curves (the dependence of extinction on wavelength) can be analysed to yield size and composition of the dust particles. A useful model proposed by Mathis, *et al.* (1977) assumes a power-law distribution in particle size with many small and few large particles. If the radius of a particle is a and C is a constant, the number of particles $N(a) = C a^{-q}$. The carbon particles are in the form of graphite with a size range 0.005 to 1 μm , and the sizes of other particles range from 0.025 to 0.25 μm . Typically, $q \approx 3.5$. Observations show that the dust distribution is not the same in all directions, and best-fit model parameters vary from region to region.

precise, easily recognised wavelength is emitted. By observing this 21-cm line and by measuring the Doppler shift owing to relative motion, the distribution of H has been mapped and found to be concentrated in spiral arms. Indeed, neutral H maps define the spiral arms of our Galaxy and of other galaxies. The presence of other elements in the ISM gas is manifested through resonance absorption lines in the spectra of stars, particularly in

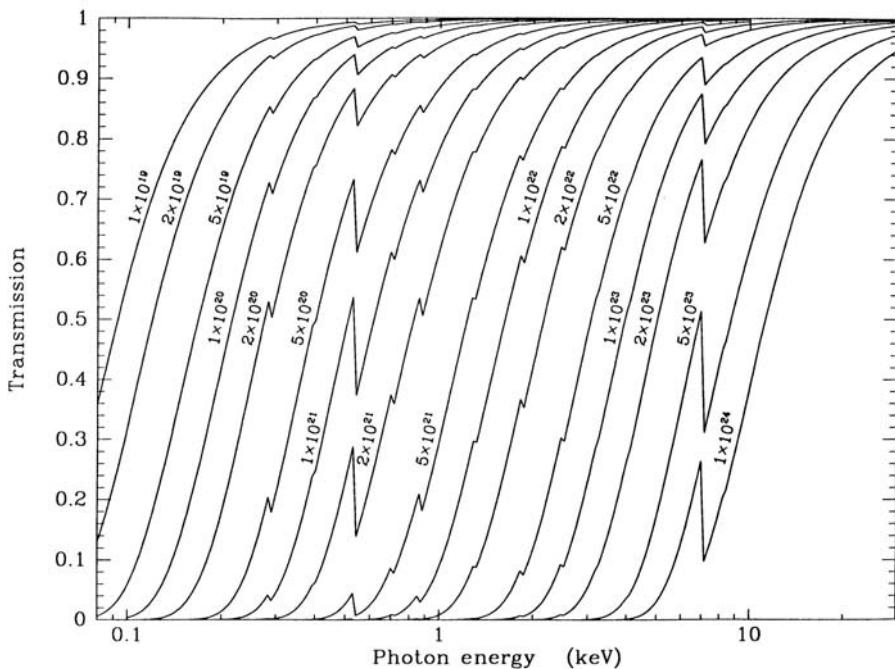


Fig 5.1 Transmission of X-rays through various amounts of interstellar gas as a function of X-ray energy. Labels on the curves give column density in atoms/cm², with column densities $< 10^{19}$ cm⁻² being essentially transparent to all X-rays > 0.1 keV. There are features because of the K edges of O and Fe at 0.5 and 7.1 keV, respectively. Transmission is 100 per cent at high energies, except for high column densities, where Compton scattering becomes important (from Seward, 2000).

the UV range, where strong lines from C and O are commonly seen. There are also molecules in interstellar space, found through radiofrequency transitions and concentrated in cold molecular clouds where the gas density is much higher (10^3 – 10^4 atoms/cm³), and molecules can exist because they are protected from the optical and UV radiation that would otherwise break them up.

X-rays are both absorbed and scattered by the ISM. The X-ray observer must cope primarily with the absorption process, but scattering can also be appreciable for distant galactic sources. Both effects are a nuisance when trying to derive source characteristics, but both also yield information about the ISM.

5.2 Absorption of X-rays by interstellar gas

The diffuse ISM comprises four components: cold neutral, warm neutral, warm ionised and hot

ionised material. The first three all absorb X-rays and can be studied by precise spectroscopic observations. The hot material is a source of diffuse emission which contributes to the soft X-ray background.

X-rays are primarily absorbed through the photoelectric effect (see Chapter 2). The transmission of the ISM is $e^{-\sigma N_H}$; σ is the photoelectric absorption coefficient which is large at low energies and decreases rapidly with increasing energy, except for abrupt increases corresponding to the threshold energies for ejection of electrons from particular atomic shells. The K-shell energies of the light elements O–Fe are in the keV range, and these are of most concern in X-ray observations. It is customary to refer to the column density of interstellar gas in units of H atoms per centimeter squared, N_H , although H itself makes little contribution to the absorption. Figure 5.1 shows calculated transmission through cold interstellar gas of varying depth. The gas is assumed to have cosmic abundance, and for each element, the K-shell

absorption feature has been approximated by an increase in absorption at a particular energy, the ‘K edge’. This absorption is present in all X-ray observations, except for Solar System targets and for the very closest stars.

For targets at ~ 2 kpc distance, the column density is $\sim 3 \times 10^{21} \text{ cm}^{-2}$, and no photons with energies below ≈ 0.4 keV are transmitted. Bright sources in the vicinity of the galactic centre, at ~ 8 kpc distance, are seen through a column of $\sim 3 \times 10^{22} \text{ cm}^{-2}$, and all X-rays with energy below ≈ 1.2 keV are absorbed. Thus, in the galactic plane survey shown in Figure 1.6, the bright sources around the galactic centre dominate the sky in the 2–10 keV range but disappear at lower energies. Other examples of absorption in the ISM are visible in Fig 8.15 and 8.18, where observed intensity falls rapidly at low energies.

A line of sight perpendicular to the galactic plane, however, passes through a column of only $\sim 3 \times 10^{20} \text{ cm}^{-2}$, and X-rays from distant extragalactic sources can be observed down to ≈ 0.2 keV energy. Note that there is often more absorption than expected from the ISM in the Milky Way. Sources in distant galaxies are also viewed through the ISM of the host galaxy, and the sources themselves often have intrinsic absorbing gas associated with e.g. accretion discs.

Although the gas column density is given in H atoms per centimeters squared, it is not the H that is responsible for most of the absorption. At 0.2 keV, 60 per cent of the photons absorbed are absorbed by He. At 0.6 keV, O accounts for 45 per cent of the absorption. Thus the observed soft cutoff for a column of $3 \times 10^{21} \text{ cm}^{-2}$, where transmission drops from 80 per cent to 20 per cent is largely a measure of the amount of interstellar O in the line of sight. Figure 5.2 shows the relative effect of principal elements in the ISM.

K edges of several absorbing elements have been directly observed in high-resolution spectra. Figure 5.3 shows an example of the O edge at ≈ 0.53 keV, as seen in the spectra of two bright accretion-powered binaries. This shows the complexity of modern data. What used to be interpreted as a simple edge is revealed to contain resonance features and edges from three ion species. Some of the OII and OIII absorption occurs close to the sources; the neutral O (OI) is in

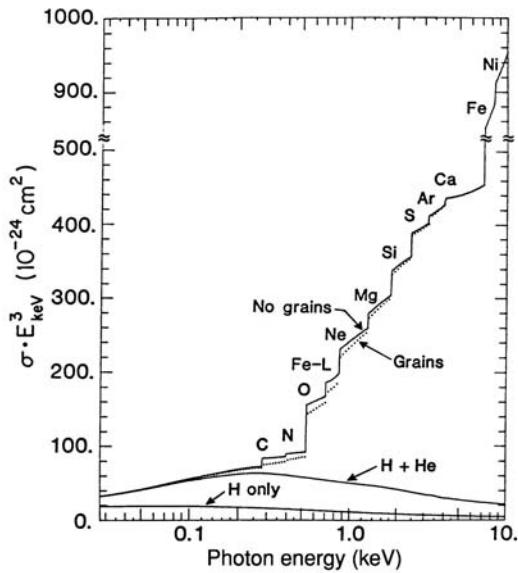


Fig 5.2 A plot showing the relative absorption owing to different elements of the ISM. At 0.2 keV, 60 per cent of the absorption is in He. At 0.6 keV, 45 per cent is in O, and at 8 keV, 45 per cent is in Fe (from Morrison & McCammon, 1983).

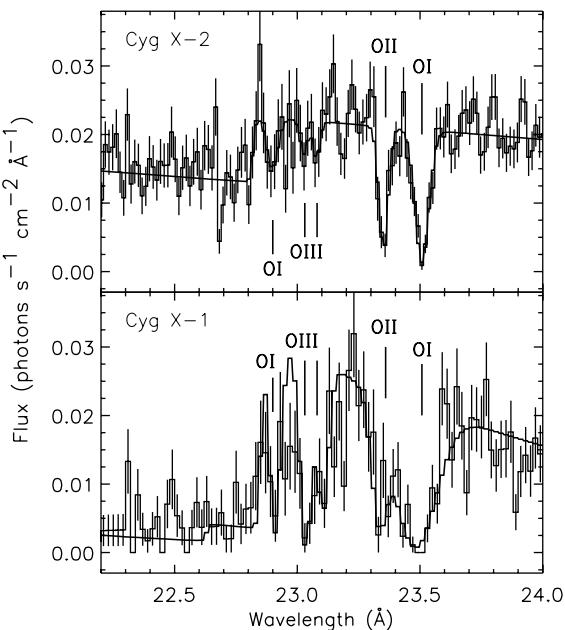


Fig 5.3 Interstellar absorption in the Chandra HETG X-ray spectrum of Cyg X-1 and Cyg X-2. Resonance absorption and edges of both neutral and ionised O are present. The K edge is clearly not a simple feature (from Juett et al., 2004).

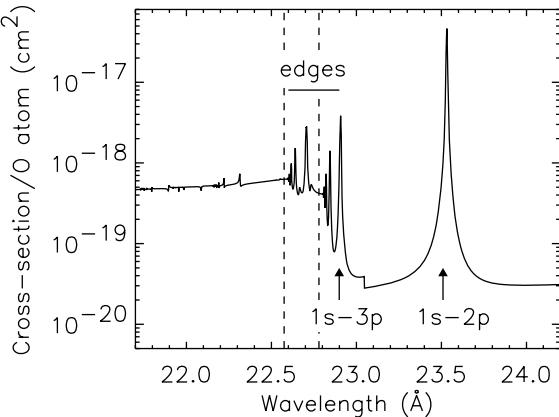


Fig 5.4 The theoretical K-edge absorption probability of neutral oxygen. Each absorption peak is from a specific electron transition from the K shell to an outer shell. In the data, 1s-2p features for OI, OII and OIII are prominent (from Juett *et al.*, 2004).

the ISM, and Fig. 5.4 shows the expected absorption from OI. This structure was first observed by Schattenburg and Canizares (1986) using the Einstein spectrometer. Figure 5.3 shows Chandra data from Juett *et al.* (2004). They conclude that along these paths the ratio of OII to OI in the ISM is ≈ 10 per cent.

5.3 | Shadows

An interstellar cloud can easily be dense enough to absorb soft X-rays. A 10-pc cloud with an average number density of 30 atoms cm^{-3} constitutes a column of 10^{21} atoms cm^{-2} , which, as can be determined from Fig. 5.1, will absorb over 80 per cent of X-rays with energy less than 0.3 keV. Such a cloud, located between observer and a uniformly emitting X-ray background, will cast a shadow. Figure 5.5 shows one of the first detections of such a shadow. The ROSAT instruments, with wide field of view, high sensitivity to soft X-rays and low background, were well suited for such observations. The figure shows absorption of diffuse background X-rays by a cloud located in the constellation Draco, about 300 pc distant. Sixty per cent of the X-rays from the 1/4-keV band are absorbed in the cloud, indicating that much of this background originates beyond the cloud.

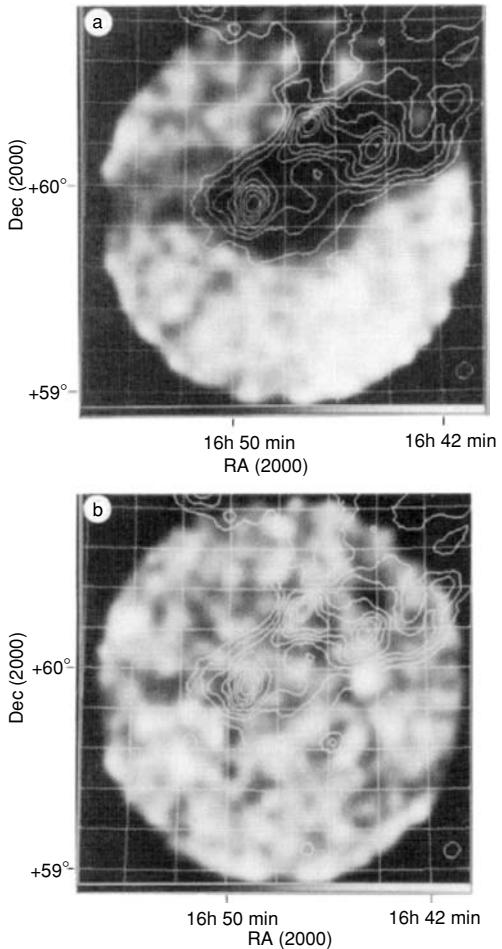


Fig 5.5 An X-ray shadow cast by the Draco Cloud. Both figures cover the full 2° diameter field of the ROSAT PSPC detector. Contours indicate the outline of the cloud as seen at $100 \mu\text{m}$ in the infrared. Point sources have been removed to show the almost uniform background. (a) The 0.13–0.28 keV energy band, where there is considerable absorption in the cloud. (b) The 0.5–2.0 keV band, where there is no shadow (from Burrows & Mendenhall, 1991).

Such data can be used in determining the origins of the soft background, which is discussed in Chapter 16.

The shells of some supernova remnants are dense enough to cast shadows. There are two cases known of overlapping remnants where the more distant remnant is shadowed by the nearer. The bright remnant IC 443 (0.75° in diameter) is behind a fainter older remnant, G189.6+3.3, discovered by Asaoka and Aschenbach (1994) and 1.5°

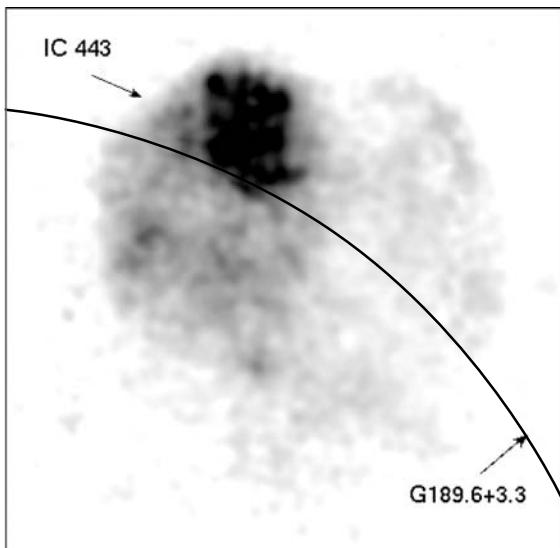


Fig 5.6 Absorption of X-rays from the SNR IC 443 by material in the limb of a larger, nearer shell-like remnant. The image of IC 443 is a smoothed mosaic of Einstein IPC data in the energy range 0.5–4 keV. The diameter of IC 443 is 45', and the X-ray structure is dominated by the shell's interaction with molecular clouds. The limb of G189.6+3.3 is indicated by a solid line, and material in the shell casts a shadow which appears as a lighter band just inside the rim. X-rays from G189.6+3.3 are too faint to show in this figure (Einstein data are from Seward, 1990; G189.6+3.3 from Asaoka & Aschenbach, 1994).

in diameter. The shell of G189.6+3.3 lies directly in front of the bright interior of IC 443 and absorbs some of the emission from IC 443. Figure 5.6 shows this. The derived column density of the shell is $2 \times 10^{21} \text{ cm}^{-2}$ of cold material. A similar situation occurs with the 8° diameter Vela remnant and the smaller, brighter remnant Puppis A.

5.4 | Absorption by intergalactic gas

Absorption of X-rays by diffuse matter between galaxies has also been observed and, for completeness, is mentioned here. Using the Chandra LETG/ACIS detector combination, absorption of the red-shifted ions O^{VII} and N^{VII} has been detected in the spectrum of the blazar Mrk 421 (Nicastro *et al.*, 2005). To get the large number of photons required, the observation was done as a

target of opportunity when Mrk 421 was in outburst, and two outbursts were observed for more than a day each. Although the column density of the absorbing material is small, the volume of space and the mass of gas are huge. This material may represent a significant fraction of the baryons in the Universe, so the cosmological implications are important, but data are difficult to obtain and interpret. In particular, the contribution from hot gas in the halo of our Galaxy is not well determined and consequently difficult to subtract.

5.5 | Scattering of X-rays by interstellar dust

Interstellar dust scatters X-rays. Ten to 30 per cent of the X-rays from sources in the vicinity of the galactic centre are scattered through angles of approximately a few arc minutes and appear as a distinct halo around each source. Look at a bright, distant light through a misty window e.g. from a car at night, and you will see a similar effect.

The scattering of X-rays by interstellar dust was predicted by Overbeck (1965) and first observed but not recognised by Toor *et al.* (1976) using a rocket-borne detector during a 1974 lunar occultation of the Crab Nebula. With the launch of Einstein in 1979, useful imaging was possible, and Dan Rolf discovered a dust halo around the bright galactic source GX 339-4. He interpreted this simply as being caused by a uniform distribution of single-sized grains and derived a grain radius of $\approx 0.06 \mu\text{m}$ (Rolf, 1983). Dust halos amounting to 5 to 10 per cent of the source intensity were subsequently observed around bright galactic sources by Catura (1983) and Mauche and Gorenstein (1986) with Einstein and by Predehl and Schmitt (1995) with ROSAT. Analysis in all cases was accomplished by showing that the measured radial extent of the source was larger than that expected from the point-spread function of the telescope mirror. It was found that the amount of scattering by dust was well correlated with the measured absorption of soft X-rays by the interstellar gas. The gas and dust are mixed together, not located in different regions. So far, halos have been found to be symmetrical in azimuth, so

Box 5.2 | Scattering from dust

For particles of radius a μm and X-rays of energy E keV, the size of the halo is $\approx 1.0(aE)^{-1}$ arcmin. The intensity of the halo is determined by the amount of dust and by the cross section for scattering, which depends on size and composition of the dust. An approximate formula for the total scattering cross section is

$$\sigma = 6.3 \times 10^{-7} (2Z/N)^2 (\rho/3)^2 a^4 E^{-2} \text{ cm}^2.$$

Z , N and ρ are atomic number, atomic weight and density of the dust particles in gm/cm^3 , respectively. Because of the E^{-2} factor, soft X-rays are scattered more strongly than hard X-rays, and the strong a^4 dependence on particle size means that large particles dominate the scattering process. Also note that for a given particle size, the halo observed at low energy will be larger than the halo observed with more energetic X-rays.

the dust is not strongly clumped in clouds. Note, however, the structure observed around Cyg X-3, discussed at the end of this chapter.

The pure dust-scattered halo was observed by ROSAT during a 1991 lunar occultation (Predehl *et al.*, 1992). Colour Plate 9 shows ROSAT images taken just before and just after GX 5-1 appeared from behind the lunar limb. When the source itself, which accounted for 70 per cent of the emission, was occulted, the mirror-scattered contribution disappeared, and the remaining signal was from only the dust-scattered halo. Although this was a fascinating observation, longer exposures are needed to measure extent and shape of dust halos.

Modern telescope mirrors are well calibrated, particularly at low energies, where dust scattering can be dominant, and the instrumental contribution can be subtracted with confidence. Figure 5.7 shows an unprocessed Chandra image of GX 13+1, a strongly absorbed accretion-powered binary 8 kpc distant. All X-rays with energy < 1.2 keV have been absorbed in the interstellar gas. A bright halo with radius $\sim 10'$ is observable surrounding the source. The telescope mirror is good enough so that the halo is well separated from the core of the point-source response

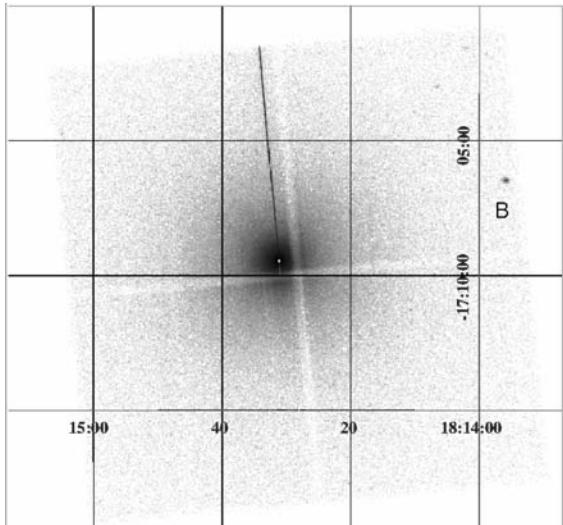


Fig 5.7 An unprocessed Chandra image showing GX13+1 surrounded by a scattering halo. The exposure was 3 hours, and most of the events are in the energy range 2–5 keV. The source itself appears as a white dot at the centre of the halo because pileup in the detector causes events in the central pixels to be rejected (an artefact peculiar to CCD detectors). The intrinsic angular extent of the point source in this image can be seen as the width of the charge-transfer streak (another CCD artefact) extending north from the source position. Gaps between the four ACIS chips appear as a cross-shaped shadow – fuzzy because of telescope dither (from Smith *et al.*, 2002).

function. The bright part of the halo arises from X-rays in the energy range 2–5 keV. Figure 5.7 also shows three artefacts peculiar to the Chandra ACIS detector which make the analysis non-trivial.

Although the Chandra X-ray mirror is state of the art, imperfections still cause small-angle scattering, and the halo seen in Colour Plate 9 arises from a mixture of mirror scattering and interstellar dust scattering. However, at low energies, dust scattering dominates. After subtraction of the mirror contribution, the shape of the dust halo can be determined at different energies. Shapes at two energies from the halo shown in Fig. 5.7 are shown in Fig 5.8. The Mathis, *et al.* (1977) model was found to fit the data fairly well, assuming a uniform distribution of dust. Figure 5.8 also shows fits calculated assuming that the dust is concentrated in three distinct regions, or clouds, between us and the source. The fits are as good as those for a uniform distribution, and only three regions are

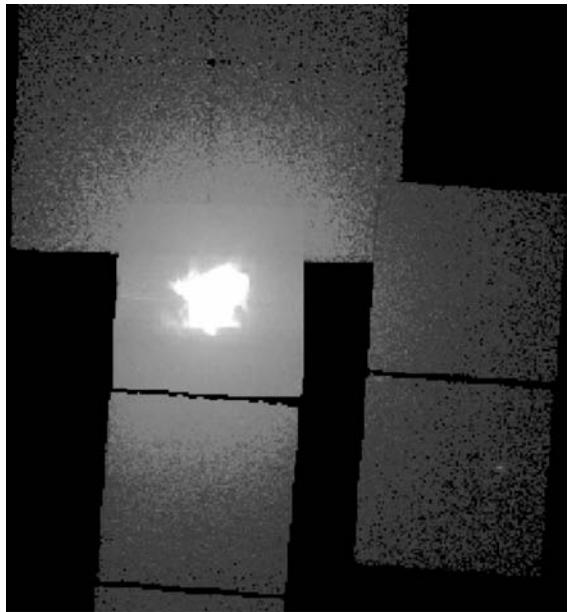
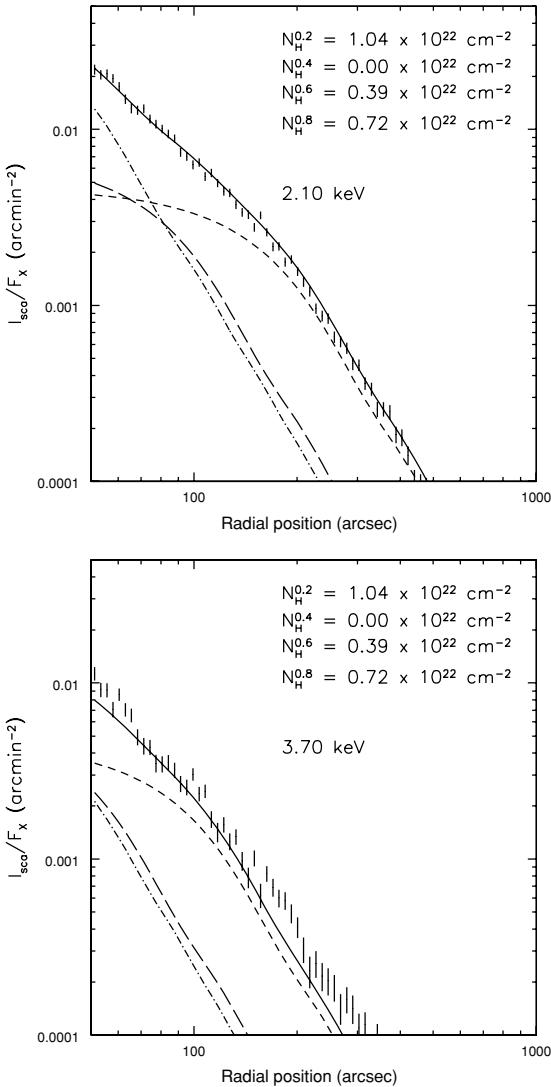


Fig 5.9 Three merged Chandra observations of the Crab Nebula and vicinity. The bright, overexposed nebula is seen embedded in a dust-scattered halo with angular extent of at least 20 arcmin. Each square is a Chandra ACIS chip $8' \times 8'$ in dimension (Fig. 9.4 shows the Crab with shorter exposure) (from Seward *et al.*, 2006).

needed. Because concentration of dust is expected in spiral arms, this result is not surprising.

Figure 5.9 shows three merged Chandra observations of the Crab Nebula. The CT streaks have been removed, and the data have been corrected for vignetting and then smoothed. The observed halo extends more than 15' from the source, is largely due to X-rays with energies ≈ 1 keV and has an integrated brightness 5 per cent that of the Crab Nebula itself. Because the absorbing column of gas is less than that of GX 13+1, the softer X-rays are not absorbed, and the observed halo is larger. Predictions of various dust models were compared with the data, and good fits required that ~ 10 per cent of the dust be concentrated at the near edge of the absorbing/scattering column.

Mitsuda *et al.* (1990), using Ginga, observed a 1987 10-min lunar eclipse of a transient source near the galactic centre. X-rays from the halo were visible for about a minute just after ingress (the sudden eclipse of the pointlike source by the limb of the Moon) and again just before egress (the

reappearance of the source). Because the instrument was a simple proportional counter, there was no contribution from instrumental scattering. The size of the halo was used to estimate a typical grain size of $0.06 \mu\text{m}$, which, because of the a^4 dependence of the cross section, should be close to the maximum grain size. The spectrum of scattered X-rays from 2 to 10 keV was measured and fit with models using different elemental compositions. It was concluded that the dust contained a large amount of Fe, although, because the signal was small, uncertainty was large.

5.5.1 Time-dependent scattering phenomena

A different aspect of dust scattering is observable when the strength of the source varies. Because the path length of scattered light is longer, scattered X-rays arrive later than those proceeding directly from the source.

5.5.1.1 γ -ray bursts

A γ -ray burst can produce a pulse of X-rays (mostly from the afterglow) which lasts a short time compared to the time required to record images. A short observation, taken after the initial burst and airglow have decayed, records both photons emitted during the exposure and photons from the initial burst which have taken a path longer than the direct path from source to detector. If there is a solitary dust cloud in the line of sight, the scattered X-rays will appear as an expanding ring around the source position. If the γ -ray burst (GRB) is very far away, the angular radius of the halo is $\theta = (2ct/d)^{1/2}$, where c is the speed of light, t is the time since the burst and d is the distance from Earth to the scattering cloud.

The first observations of this phenomenon were made by Vaughan, *et al.* (2004) using an XMM observation of GRB 031203. A ring from GRB 050724 using a small telescope aboard SWIFT has also been imaged and is shown in Colour Plate 10. The measured rate of expansion of the GRB 050724 ring determined that the scattering cloud was at a distance of 139 ± 9 pc. Note the accuracy with which the distance to the cloud was measured and that the expansion of the halo was fairly rapid. This observation was accomplished in the first 3 hours after the burst.

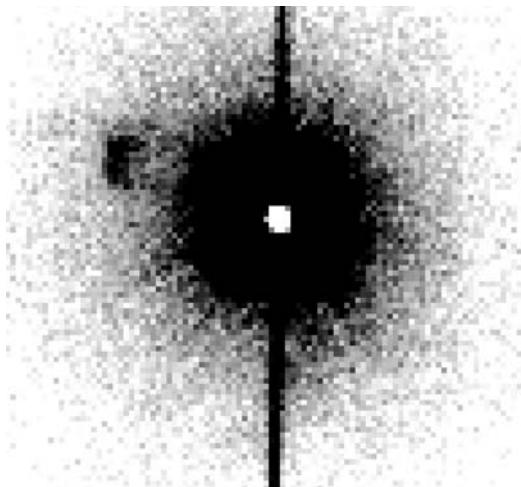


Fig 5.10 A Chandra observation of Cyg X-3 showing structure in the observed halo, about half of which is due to scattering from dust in the ISM. The dust cloud which produces the bright spot $\approx 16''$ east of the source lies at three-fourths of the distance to Cyg X-3. The field shown is small, only $45''$ square. The bright CT streak, the central hole and half the halo intensity are instrumental (from McCollough, 2009).

5.5.1.2 Cyg X-3

Figure 5.10 shows a Chandra observation of Cyg X-3, a bright, strongly absorbed galactic source. The instrument/dust-scattering halo is bright and contains a distinct structure $\approx 16''$ east of the source. This observation by Mike McCollough was the first indication of strong azimuthal structure in an X-ray dust halo. It maps a localised concentration, or cloud, of dust. This $3'' \times 5''$ spot has surface brightness 1.5–2 times the average value at this radius, and the total intensity is $\approx 10^{-4}$ that of Cyg X-3. Geometry indicates that the cloud must be at least 2.7 light years from Cyg X-3, but the timing enables us to determine a rather accurate location.

X-rays from Cyg X-3 are strongly modulated with the orbital period of 4.8 hours. The halo brightness, being half instrumental, also varies with this period and follows the source intensity exactly. The brightness of the spot also follows the sinusoidal source variation, but with a phase lag of 0.55, which corresponds to a 2.6-hour delay. Because light travel time depends on scattering angle, this phase delay indicates that the distance to the cloud is 0.77 the distance to the source, or

about 7 kpc. The size of the cloud then is ≈ 0.10 by 0.17 pc (McCollough, 2009). Accurate subtraction of the time-varying halo background was a non-trivial operation and vital to this result.

5.5.1.3 4U 1538-52

The source 4U 1538-52 is another example of a time-dependent phenomenon. This source is a binary with period 3.7 days and a 12-hour eclipse. Clark (2004) used Chandra to image the halo of 4U 1538-52 starting just before ingress and lasting until mid-eclipse. Profiles of the observed halo were measured at three times. Each radial profile had a distinct three-peak shape implying dust concentration at three distances. For each distinct dust cloud the small-angle scattering disappeared soon after the eclipse and the observed scattering halo lost its bright centre. By observing the halo shape at different times and by fitting with models, Clark derived distances of the dust concentrations. The situation was not as clean as that of a single burst of X-rays, but the sharp cutoff of emission provided an observable signature from which definite information was derived.

The preceding examples illustrate that effects of absorption and scattering are present in all X-ray observations. Absorption is usually measured, and results corrected for this effect. Scattering is usually ignored. This is reasonable for most extragalactic observations with little absorption. For strongly absorbed galactic sources, the dust should be taken into account.

References

- Asaoka, I. & B. Aschenbach (1994) An X-ray study of IC 443 and the discovery of a new supernova remnant by ROSAT, *A&A* **284**, 573.
- Burrows, D. N. & J. A. Mendenhall (1991) Soft X-ray shadowing by the Draco cloud, *Nature* **351**, 629.
- Catura, R. C. (1983) Evidence for X-ray scattering by interstellar dust, *ApJ* **275**, 645.
- Clark, G. W. (2004) Chandra observations and Monte Carlo simulations of the grain-scattered halo of the binary X-ray pulsar 4U 1538-52, *ApJ* **610**, 956.
- Juett, A. M., N. S. Schulz & D. Chakrabarty (2004) High-resolution X-ray spectroscopy of the Interstellar Medium: Structure at the oxygen absorption edge, *ApJ* **612**, 308.
- Mathis, J. S., W. Rumpl & K. H. Nordsieck (1977) The size distribution of interstellar grains, *ApJ* **217**, 425.
- Mauche, C. W. & P. Gorenstein (1986) Measurements of X-ray scattering from interstellar grains, *ApJ* **302**, 371.
- McCollough, M. (2009) The true nature of Cygnus X-3's 'little' friend, paper presented at Chandra's first Decade of Discovery, 22–25 September, Boston.
- Mitsuda, K., et al. (1990) Dust-grain scattering of X-rays observed during the lunar occultation of a transient source near the galactic center, *ApJ* **353**, 480.
- Morrison, R. & D. McCammon (1983) Interstellar photoelectric absorption cross sections, 0.03–10 keV, *ApJ* **270**, 119.
- Nicastro, F., et al. (2005) Chandra detection of the first X-ray forest along the line of sight to Markarian 421, *ApJ* **629**, 700.
- Overbeck, J. (1965) Small-angle scattering of celestial X-rays by interstellar grains, *ApJ* **141**, 864.
- Predehl, P. & J. H. M. M. Schmitt (1995) X-raying the Interstellar Medium: ROSAT observations of dust scattering halos, *A&A* **293**, 889.
- Predehl, P., et al. (1992) A lunar occultation of the dust-scattering halo around GX 5-1 observed with ROSAT, *Science* **257**, 935.
- Rolf, D. P. (1983) Evidence for the detection of X-ray scattering from interstellar dust grains, *Nature* **302**, 46.
- Schattenburg, M. L. & C. R. Canizares (1986) High-resolution X-ray spectroscopy of the Crab Nebula and the oxygen abundance of the Interstellar Medium, *ApJ* **301**, 759.
- Seward, F. D. (1990) Einstein observations of galactic supernova remnants, *ApJ Suppl* **73**, 781.
- Seward, F. D. (2000) *Astrophysical Quantities*, edited by A. Cox 183, New York: AIP Press.
- Seward, F. D., P. Gorenstein & R. Smith (2006) Chandra observations of the X-ray halo around the Crab Nebula, *ApJ* **636**, 873.
- Smith, R. K., R. J. Edgar & R. A. Shafer (2002) The X-ray halo of GX 13+1, *ApJ* **581**, 562.
- Toor, A., T. Palmieri & F. Seward (1976) Evidence for 'thermal' X-rays from the Crab Nebula, *ApJ* **207**, 96.
- Vaughan, S., et al. (2004) The discovery of an evolving dust-scattered X-ray halo around GRB 031203, *ApJ* **603**, L5.
- Vaughan, S., et al. (2006) The dust-scattered X-ray halo around SWIFT GRB 050724, *ApJ* **639**, 323.

Chapter 6

Active stellar coronae

6.1 | The Sun

The Sun is close and has been studied intensively. It radiates strongly from radio- to X-ray frequencies and, because of solar-terrestrial effects, has been monitored by an armada of spacecraft for 50 years. There were the OSO spacecraft (which also observed other cosmic sources) (1962–1978), Skylab (1973), Solar Max (1980–1989), Yohkoh (1991–2001), SOHO (1995–), TRACE (1998), and Hinode (2006). Solar X-ray emission is now continuously measured by a series of GOES spacecraft, and current data are available online almost instantaneously (NOAA/SWPC, 2009a). In this section we show only a few observations which illustrate things to keep in mind when considering the emission of other stars. The data are spectacular, and we regret not having room to include more. For a more thorough overview of solar observations and theory, there is an excellent book by Golub and Pasachoff (1997). Movies of EUV and X-ray images of the Sun can be viewed on several websites (e.g. TRACE, 2009; XRT, 2009).

6.1.1 An historical puzzle

Why should there be detectable X-rays from the Sun at all? Certainly not on the basis of its everyday visible appearance. The optical spectrum of the Sun can be represented quite well by a simple blackbody at a temperature of about 6000 K. Such an object should produce no detectable X-ray flux, whereas the amount actually seen implies

the presence of material at a temperature of at least 1 million degrees!

In fact, the X-ray observations of the Sun represented the last pieces in a puzzle about the nature of the Sun's corona that began in the middle of the nineteenth century. The corona is only visible to us on Earth during a total eclipse by the Moon. When the first spectroscopic observations were made of the corona during an eclipse more than 100 years ago, the results were puzzling. Spectral lines, never seen in the laboratory, indicated the presence of an unknown element in the corona. The new element was dubbed *coronium*.

Further work by Lyot in the 1930s found very broad emission lines in the coronal spectrum. If this were a velocity effect, then it would require a temperature of more than 1 million degrees. By then the emission lines of coronium had been identified as owing to nothing more exotic than iron and other heavy elements! They had not been identified earlier because these coronal elements had been stripped of a large number of their outer electrons, leaving them with a net positive charge. This requires temperatures of 1–2 million degrees.

The existence of a million degree solar corona was then accepted without question, and the detection of thermal X-rays removed all doubt. But what supports the corona and provides it with its energy? It does not make sense for the temperature to increase above the Sun's surface; rather, one would expect it to decrease. There was no obvious mechanism by which the material in the corona could be heated. And without a heat source, it would very rapidly cool (on a timescale of tens of minutes).

Box 6.1 | The nature of the Sun

The Sun is a sphere of hot gas which, although rather strange to say so in an astronomy book, is truly huge. The mass ($1 M_{\odot} = 2 \times 10^{33}$ g) accounts for 99.8 per cent of that of the solar system, and its visual radius is $R_{\odot} = 7 \times 10^5$ km. By weight the outer part of the Sun is 71 per cent H, 26.5 per

cent He and 2.5 per cent other elements. Density and temperature are extremely high at the centre and decrease rapidly with increasing radius. At the ‘surface’ and beyond it is more complex. Table 6.1 gives some numbers.

Table 6.1 | Layers of the Sun.

Layer	Radial position	Density	Temperature (K)
Centre	$R = 0$	160 gm cm^{-3}	$1.5 \times 10^7 \text{ K}$
Halfway out	$R = 0.5 R_{\odot}$	1 gm cm^{-3} (like water)	$3 \times 10^6 \text{ K}$
Photosphere	$R = 1.00 R_{\odot}$	$\sim 10^{17} \text{ cm}^{-3}$	6000 K
Chromosphere	$R = 1.001 R_{\odot}$	$\sim 10^{14} \text{ cm}^{-3}$	5600 K
Chromosphere	$R = 1.003 R_{\odot}$	$\sim 10^{11} \text{ cm}^{-3}$	8000 K
Corona	$R = 1.004 R_{\odot}$	$\sim 10^9 \text{ cm}^{-3}$	$\sim 10^6 \text{ K}$
Corona	$R = 1.15 R_{\odot}$	$\sim 10^8 \text{ cm}^{-3}$	$\sim 1.5 \times 10^6 \text{ K}$

We see the *photosphere*, a visually opaque layer ~ 100 km thick and with temperature of 6000 K. This radiates a blackbody spectrum with total luminosity $L_{\odot} = 4 \times 10^{33}$ ergs s^{-1} . Above this is a 2000-km-thick layer, the *chromosphere*, which is slightly cooler. Here absorption produces strong, dark spectral lines (the Fraunhofer lines), and there is also emission characteristic of the material therein. In this layer the temperature does not change much, but the density falls a factor of 10^7 . There is then an abrupt decrease in density and an increase in temperature to $\sim 10^6$ K. This is the start of the outer atmosphere, the *corona*, and the hot gas in the corona is a strong source of thermal X-rays. White light from the corona can be seen during a solar eclipse which, when the Sun is active, can extend out to several R_{\odot} . At some point, material from the hot, expanding corona becomes the solar wind.

Every layer of the Sun is pulled towards the centre by gravity but supported by the flow of energy from the interior. Nuclear energy released in the core ($R < 0.3 R_{\odot}$) diffuses outwards by radiative transfer (photons colliding with particles) until $R \approx 0.8 R_{\odot}$, where, as shown in Fig. 6.1, the process

becomes convective. Doppler shifts of radiation from rising and falling convected material are observed at the surface.

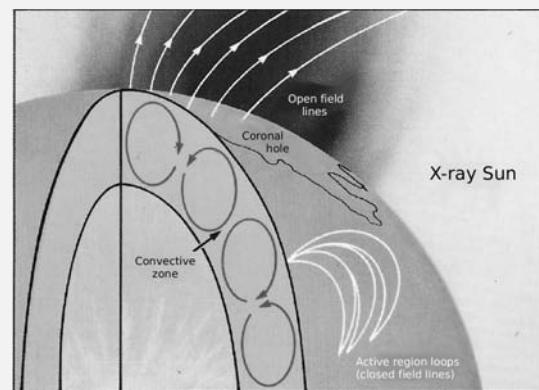


Fig 6.1 | A cross section of the Sun showing convection zone and magnetic fields in the corona.

The Sun rotates, and because it is not solid, different parts rotate at different speeds. At the equator the sidereal rotation period is ≈ 25 days; at a latitude of 45° the period is ≈ 28 days. Solar oscillations indicate that the rotation period also varies with depth. Interior layers rotate faster

down to $R \approx 0.9R_{\odot}$ then more slowly. (Turbulence in the convection zone creates pressure waves which travel through the solar interior. The Sun ‘rings’ at characteristic frequencies, which depend on interior structure. A typical frequency is 1 cycle/5 min, and the waves are detected via Doppler shifts at the surface.)

A magnetic field is generated in the interior, probably by differential rotation at the diffusive-conductive zone boundary. This field is carried to the surface by convected material and plays a dominant role in heating the corona and in the appearance of the surface.

The surface usually has a number of large, dark spots, sunspots, somewhat cooler than the surface and seen clearly in the light of H α emission. The number visible varies from 0 to ~ 200 , as shown in Fig. 6.2, and follows an 11-year cycle. The magnetic field at the surface varies continuously but is generally in the range 0–100 G. The field in a large sunspot can be as high as 2–4 kG. The solar area

covered by sunspots is never large. Even at solar maximum, the fraction of the solar surface covered is < 0.1 per cent. Sunspots, however, are the locations of active regions and solar flares, where solar X-ray emission is brightest.

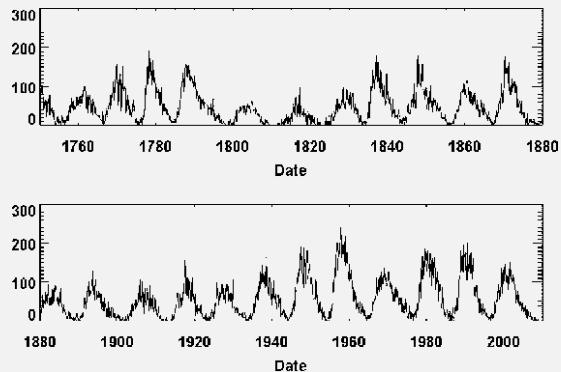


Fig 6.2 Two hundred sixty years of sunspot data. The 11-year solar cycle has been well recorded over this interval. Note that the maximum level of activity varies from cycle to cycle (data from SIDC Belgium, 2009).

A detailed photo of the Sun’s surface shows a granular structure which we know to be from cells of turbulent motion. These cells are the key to two models that explained the heating of the solar corona: acoustic and magnetic heating. Both rely on energy provided by turbulence in the Sun’s convection zone just below the surface. Shock waves generated by this turbulence were thought capable of propagation through the surface and into the corona. But the favoured model now is one in which the turbulence helps amplify magnetic fields at the surface, creating loops along which electric currents can pass, thereby heating the corona. Indeed, modern images show that magnetic fields appear to determine coronal structure.

6.1.2 Structure of the corona

Although we have referred to the corona as the outer layer of the Sun, it is not that simple. It is very much a three-dimensional structure of complex and varying geometry. The gas is turbulent and confined by magnetic fields in struc-

tures which extend over large distances and which change dramatically in a few minutes. The movies of this activity are spectacular (TRACE, 2009). There is also long-lived structure that persists over several solar rotations.

Figure 6.3 shows an X-ray picture of the entire Sun. Hot gas in the corona shows as structured diffuse emission covering most of the surface. There are also two large areas of very weak emission which are called *coronal holes*. Here the magnetic field lines do not reconnect with the Sun but are carried into space by the solar wind, and when Earth passes through one of these regions, we observe an increase in the solar-wind velocity. The surface is flecked with *X-ray bright points*, tiny bipolar areas of emerging magnetic flux which have lifetimes of only hours to days. Holes and bright points are almost always present as features of the quiet Sun. When the Sun is *active*, X-ray emission is dominated by active regions associated with the sunspots. These comprise bright loop structures connecting regions of opposed magnetic polarity. It is common for loops to connect the outer region (*penumbra*) of one sunspot with

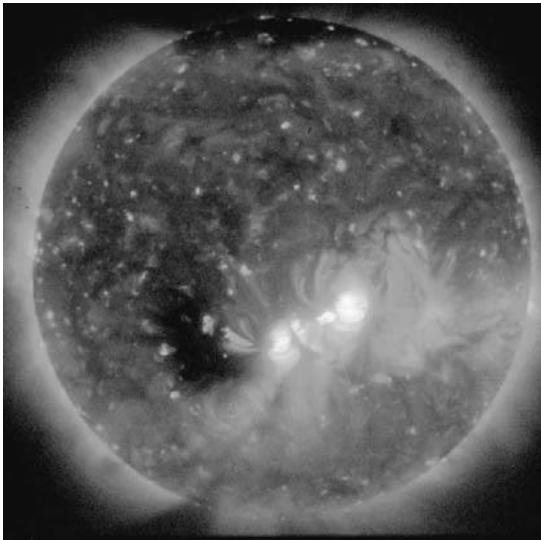


Fig 6.3 Coronal X-rays in the band 0.5–1 keV recorded by the XRT instrument on Hinode in December 2006. The brightest emission is from an active region. The quiet area is covered with diffuse emission, except for the dark coronal holes at the north pole and just left of the active region. The surface is also flecked with X-ray bright points (Hinode is a Japanese mission developed and launched by ISAS/JAXA with NAOJ as domestic partner and NASA and STFC [UK] as international partners; image from JAXA 2007).

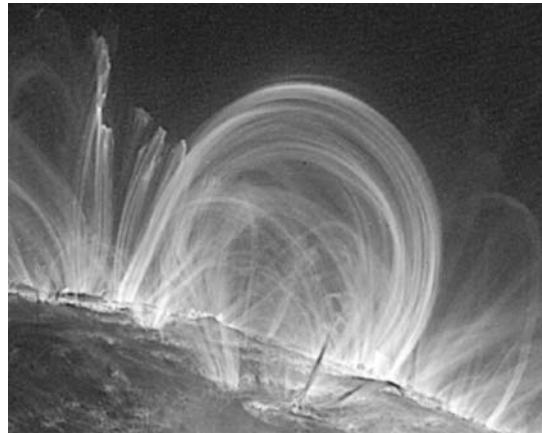


Fig 6.4 An image from TRACE showing loops at the limb of the Sun taken 6 November 1999. This is an EUV picture at 171 Å (0.075 keV), a little below the soft X-ray range. This is an emission line of Fe X which indicates (Fig. 2.2) material at temperatures of $1\text{--}2 \times 10^6$ K (image from Stanford-Lockheed Institute for Space Research, NASA/SLISR, 2009).

that of another. X-rays from an active region are bright in Fig. 6.3.

X-ray emission from the corona is thermal. The spectrum is soft and there are many spectral lines. In general, at solar minimum, when the Sun is quiet, the X-ray luminosity above 0.1 keV is $\sim 0.4 \times 10^{27}$ ergs s^{-1} , only about 10^{-7} of the optical luminosity. At solar max, there are more sunspots and active regions, and the quiet Sun L_x is $\sim 4 \times 10^{27}$ ergs s^{-1} , an order of magnitude higher. A large flare can radiate $\sim 50 \times 10^{27}$ ergs s^{-1} in the 0.2–2 keV band where the ROSAT survey was done. For the quiet Sun, $L_x/L_{\text{bol}} = 10^{-7}\text{--}10^{-6}$.

6.1.3 Solar flares

The strong fields around sunspots form complex, looplike structures (Fig. 6.4), and much energy is stored in these fields and the confined material. Often a sudden release of energy occurs, which heats the gas, accelerates particles and ejects material from the corona. This ‘explosion’ on the Sun is called a *solar flare*. The

mechanisms of energy storage and release are not well understood, but radiation from flares has been well observed. The total energy released varies from the threshold of detection up to $\sim 10^{33}$ ergs. Generally, from a large flare, a 1-min pulse of high-energy X- (or even γ) rays is followed by a ~ 20 -min burst of soft X-ray/UV/optical/radio radiation. If the orientation is favourable, particles thrown from the Sun arrive at the Earth minutes to days after the flare. A large flare can produce aurorae, ionisation of the upper atmosphere, a *magnetic storm* and a flux of high-energy particles – *solar cosmic rays*. The observation of nuclear γ rays from the Sun during flares shows that protons are sometimes accelerated to high energies.

Figure 6.5 shows X-ray emission from several solar flares. These data came from the GOES spacecraft, which monitors solar emission from geosynchronous orbit. X-ray flares are classified by letter, indicating flux order of magnitude, and number, indicating multiplier. A C flare is small, an M flare is medium and an X flare is large. As illustrated in Fig. 6.5, a flare with $1\text{--}8$ Å peak flux in the range $10^{-5}\text{--}10^{-4}$ W m $^{-2}$ ($10^{-2}\text{--}10^{-1}$ erg cm $^{-2}$ s $^{-1}$) is an M flare. An X3 flare, for example, would have peak flux of 0.3 ergs cm $^{-2}$ s $^{-1}$ in the band 1.5–12 keV.

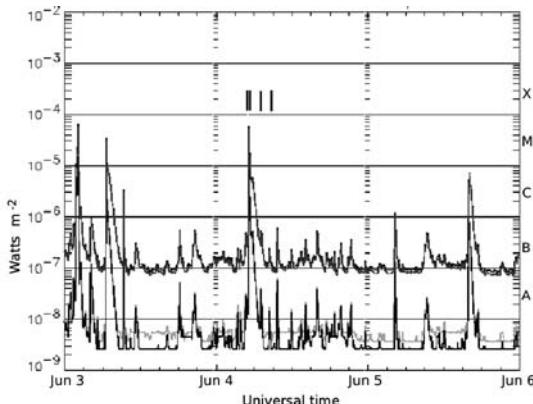


Fig 6.5 Solar X-ray emission showing moderately bright solar flares during 3–5 June 2007. The GOES satellite provides continuous monitoring of solar X-ray emission. These data are taken in two bands, (top) 1.5–12 keV and (bottom) 3–25 keV. The flare shown in Fig. 6.6 is the one occurring 4 June 0510 universal time (UT). Vertical lines mark times of the four exposures shown in Fig. 6.6. The letters on the right vertical scale show an intensity classification scheme. This flare was Class M8 (NOAA/SWPC, 2009b).

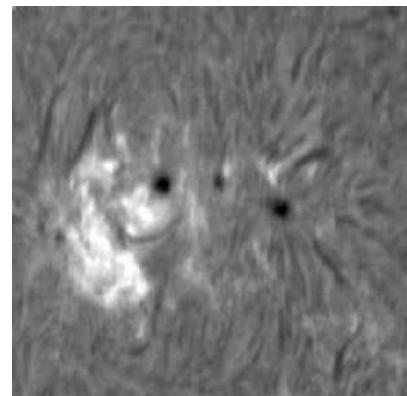


Fig 6.7 An H α picture of active region 10960 taken 9 hours after the flare shown in Fig. 6.6. The field is also 4' square and is dominated by two large sunspots (from Kanzelhoehe Solar Observatory via BBSO/NJIT, 2007).

Figure 6.6 shows the geometry of X-ray emission from an active region (AR 10960) just before and just after the M flare shown in Fig. 6.5. Figure 6.7 shows the associated group of sunspots. The most intense X-ray emission comes from a small volume, and on this occasion, the detector was saturated. The filter used for this observation limited the band pass to a narrow range $\approx 1\text{--}1.5$ keV. Temperature of the radiating gas was $\approx 10^7$ K, much hotter than the material visible in

Fig. 6.4. The flare occurs at the foot of a bright loop seen just before onset. In the third panel, an array of loops with hot material has been heated or produced by the flare. This had faded when the fourth picture was obtained an hour later. In this case, there were X-rays from material held by the field in loops high above the sunspots and from accelerated particles striking the surface.

A simple model for flares assumes a magnetic loop filled with coronal plasma. Energy is released at the top of the loop by magnetic reconnection. Electrons are accelerated, travel down the field lines and heat material at the surface. This is the *impulsive* part of the flare. The hot material then expands up into the loop, where it radiates and

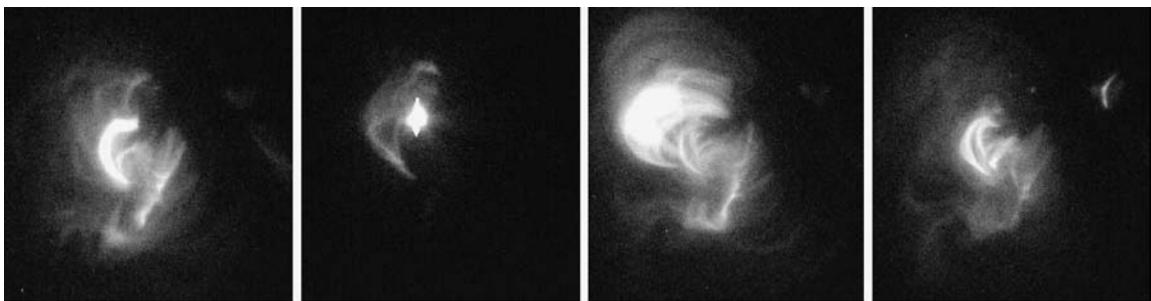


Fig 6.6 X-rays from a moderate-sized solar flare on 4 June 2007. Data are from the XRT instrument on Hinode. The field is 4' square ($\sim 10\%$ of the 30' solar diameter), and the exposure times are 33 s, except for the second frame, which is a short exposure of 0.13 s. The left panel was taken just before the flare, the next panel shows the start of the flare and the two right panels are post-flare. Maximum emission, which saturates the instrument, is not shown. Duration of this sequence is 3 hours 45 minutes (image courtesy of K. Reeve; SAO, 2009).

cools. The X-ray luminosity is determined by temperature and density of the plasma and by the size of the loop. This model has been invoked in the analysis of flares on other stars which are considerably more luminous than solar flares. Simple X-ray observations can determine luminosity, temperature, rise and decay times of a flare. High-resolution spectra can give density. We note, with caution, that individual solar flares have large differences and that flare physics is an active, ongoing area of research (meaning much is still unknown).

6.2 The Einstein stellar X-ray survey

The Einstein Observatory enabled a search for X-ray emission from other stars at a level comparable to that from the Sun (Vaiana *et al.*, 1981). The Sun's quiescent X-ray luminosity is only about 10^{27} ergs s $^{-1}$, which is 10^{10} times fainter than Sco X-1! Nevertheless, the huge gain in sensitivity afforded by Einstein's imaging telescope enabled literally hundreds of ordinary stellar sources to be found. The ordinary stars detected in this survey covered a wide range of spectral types in the Hertzsprung-Russell (HR) diagram (see Box 6.2).

Figure 6.8 shows the HR diagram with the result of the Einstein survey overlaid. This is basically a plot of a star's optical brightness against its temperature. The *main sequence* is the s curve running from upper left (the location of bright, hot stars) to lower right (the location of dim, cool stars). The important point here is that stars of virtually all spectral types were found to emit X-rays, and most were stronger X-ray sources than the Sun. Also, within any given spectral type, large fluctuations in X-ray output occurred from star to star. Both these points showed that the acoustic coronal heating model was inadequate. This is because the surface turbulence is caused by the convection zone just below the stellar surface, and because stars earlier (i.e. hotter) than about F5 have no convection zone, they would not be expected to generate any acoustic flux for a corona. Also, stars of the same spectral type

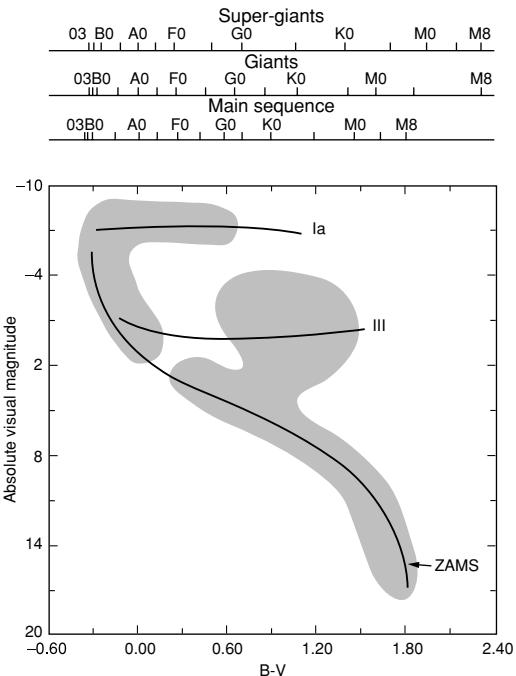


Fig 6.8 The HR diagram with results of the Einstein survey overlaid. Stars detected as soft X-ray sources fall in the grey areas. The ZAMS curve is the location of the Zero Age Main Sequence stars. Curve III shows the location of giant stars and Ia, the location of supergiants. The spectral type depends on the luminosity class and is shown above the diagram (reprinted by permission of ARAA ©Annual Reviews; from Rosner *et al.*, 1985).

should have identical convection zones and would be expected to have roughly the same coronal X-ray flux. Neither expectation was borne out by this survey.

A surprising result that came out of the Einstein survey is illustrated by observations of two well-known nearby stars. Our nearest stellar neighbour, α Centauri, is a double star, α Cen A,B, which is gravitationally connected to a dim red companion, α Cen C (called Proxima Centauri). Of particular interest, α Cen A is almost identical to the Sun. A second nearby triple is 40 Eridani, a K star orbited by an M-star/white dwarf pair. Table 6.2 lists the Einstein results. The two K stars, both cooler than the Sun and one-third as bright visually, are stronger X-ray sources than the sunlike α Cen A. The faint M stars, 30–100 times less luminous visually than the Sun, are even

Table 6.2 | X-ray emission from α Centauri and 40 Eridani.

Star	V (magnitude)	M_V (magnitude)	Spectral type	L_x (ergs s^{-1})
Sun	-26.8	4.8	G2 V	$\sim 10^{27}$
α Cen A	-0.0	4.3	G2 V	1×10^{27}
α Cen B	1.3	5.7	K5 V	3×10^{27}
Proxima Cen	11.0	15.5	M5	2×10^{27}
40 Eri A	4.4	6.0	K1 V	2×10^{27}
40 Eri B	9.5	11.1	DA	-
40 Eri C	11.2	12.7	M4	6×10^{27}

brighter in X-rays, somehow supporting extremely X-ray-luminous coronae. The ratio of the stellar brightnesses is completely reversed, and this was not expected.

After 2 years of operation, Einstein observations showed that all dwarf stars from M5 V to F5 V were X-ray emitters with luminosities ranging from the lower limit of the survey, $\approx 10^{26}$ ergs s^{-1} , to $\approx 10^{31}$ ergs s^{-1} . Many stars were observed to vary during the observations, with flaring on some. Spectra were thermal with average temperatures in the range 10^6 to 3×10^7 K. A single temperature fit, however, was not adequate for many stars for which good counting statistics were obtained.

All observed stars earlier than type B5 were X-ray sources, and the X-ray luminosity was $\sim 10^{-7}$ of the bolometric luminosity. Thus emission from some O stars was above 10^{33} ergs s^{-1} (see Chapter 7). Variability of early-type star emission was rare.

There was a narrow gap on the main sequence where X-ray emission was not detected. Furthermore, late-type giants and supergiants were also very weak emitters.

Young, pre-main-sequence stars were bright and strongly variable. Nevertheless, X-ray emission in this group decreased with age. This has been demonstrated by observations of star clusters. It is believed that all stars in a given cluster were formed at the same time, and the ages of clusters are fairly well known from numerical models of stellar evolution. We can therefore see how the X-ray emission from various clusters depends on age. You may skip ahead to Fig. 6.28, which shows X-ray luminosities of G-, K-, and M-type stars in five young clusters. The average luminosity is

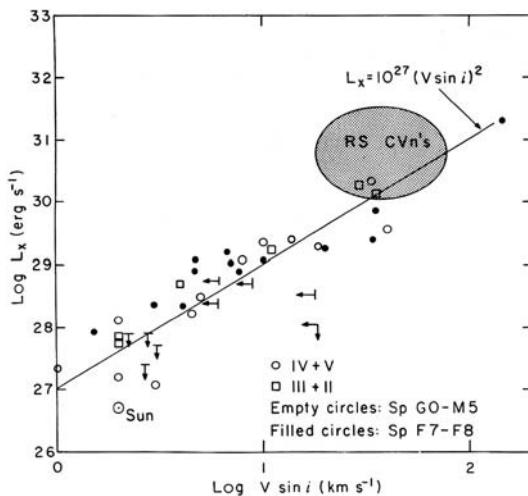


Fig 6.9 Coronal X-ray luminosity vs rotation rate. Stars which rotate faster have higher luminosities. Data are from Einstein and EXOSAT. The correlation is strong even though stars are of different spectral types (reprinted by permission of ARAA © Annual Reviews; from Rosner et al., 1985).

highest in the Orion association and decreases with age until we reach the Sun (same age as field stars), spanning a factor of 10^4 in X-ray luminosity.

There was a correlation between rotation rate of single stars and strength of X-ray emission. Figure 6.9 shows some Einstein data for stars of type M, G and F. Faster rotation produces more X-rays. Because rotation drives the dynamo which produces magnetic fields in these convective stars, this is not surprising.

Emission from close binary systems such as RS CVn and W UMa stars was considerably stronger than that expected from the sum of the two component stars.

Box 6.2 | Spectral classification and the HR diagram

Around the turn of this century, it was noted by the two astronomers E. Hertzsprung and H. N. Russell that the colour of a star was closely related to its intrinsic optical brightness. This latter quantity is usually quoted as absolute magnitude (or M_v), which is simply the apparent magnitude the star would have at a distance of 10 pc. The colour is defined as the difference in magnitude of the star in blue and visual (yellow) light (or B-V). Most stars are found to lie on the main sequence, as shown in Fig. 6.8. This figure, a plot of magnitude versus colour, is known as the HR (Hertzsprung–Russell) diagram. It is equivalent to a plot of luminosity versus temperature. When stars are formed from collapsing clouds of interstellar gas, they settle

onto the main sequence, where they spend the greatest fraction of their lives. Their position on the main sequence is determined almost entirely by their mass; the heaviest stars are the hottest (upper left of the main sequence) and the lightest stars are the coolest (lower right).

Treating stars as blackbodies, their colour is a direct indicator of temperature – blue is hot and red is cool. This is also borne out by examining the spectra of these stars, and spectral types are also shown along the x axis. Stars are principally classified by their surface temperature as derived from optical spectra. Some values are listed in Table 6.3.

Table 6.3 Physical characteristics of main sequence stars.

Spectral type	Mass M_\odot	Radius R_\odot	Surface temperature (K)	Prominent spectral absorption lines
O5	40	18	40 000	He II
B5	6.5	3.8	15 500	He I, H
A5	2.1	1.7	8500	H, Ca II
F5	1.3	1.2	6500	Ca II, H
G5	0.93	0.93	5520	Ca II, Fe, H
K5	0.69	0.74	4130	CH, CN
M5	0.21	0.32	2800	TiO

Energy leaving the star must pass through the outermost surface layer of gas. Atoms or ions in this gas absorb light at discrete energies which depend on the number of bound electrons (or the ionisation state) of the various atoms. The higher the temperature, the fewer the number of electrons that can be bound to the nucleus of the atom. Thus the particular lines observed identify the ionisation state, and the temperature can be calculated.

The ionisation of any atom is a well-known function of temperature. A cool gas, for example, can contain atomic hydrogen. Above a temperature of about 15 000 K, collisions between atoms in the gas are energetic enough to remove the bound electrons from hydrogen atoms. At higher

temperatures, hydrogen atoms are rare, and there are therefore no strong hydrogen absorption lines in the spectra of hot stars. (Astronomers refer to neutral hydrogen as HI. Ionised hydrogen is HII. Similarly neutral iron is Fe I, iron with one electron removed is Fe II. Iron with two electrons removed is Fe III, and so on.)

The hottest known stars are the O stars with spectra showing an absorption line of singly ionised helium (He II). Spectra of A stars show lines of neutral hydrogen (HI), and the coolest stars, the M stars, have temperatures low enough that molecules such as TiO can exist, giving spectra that are characterised by strong molecular absorption bands. The complete main sequence of spectral types is OBAFGKM, with temperature decreasing

from O to M. (This strange ordering of the letters is an accident. The original order of stellar spectra was based on the observed strength of the hydrogen lines. They are strongest in A-type stars, next strongest in B-type and so on. Unfortunately, the relationship with temperature is not straightforward.)

It is useful to subdivide each class into subclasses. Thus the hottest A stars are of class A0 and the coolest are A9. The Sun, with a temperature of 5800 K and strong lines of H I, Ca II and Fe II, is of spectral type G2. Table 6.3 lists characteristics of stars in the mid-range of each spectral class. Stars are further divided into luminosity classes. Table 6.3 defines the main sequence and refers to the most common stars, namely those of luminosity class V (sometimes called *dwarf stars*). The giant

stars, which have tenuous, greatly extended outer layers, are luminosity class III. The largest stars, the supergiants, are luminosity class I. The intermediate classes, II and IV, are bright giants and subgiants, respectively. The luminosity classes I to IV indicate that the star is evolved and has moved away from the main sequence. This is the process of stellar ageing, which occurs when the star exhausts its central source of nuclear fuel, causing the core to shrink and the outer atmosphere to expand.

A giant or supergiant star is not only larger but more massive than main sequence stars of the same spectral type. The Sun is a G2V star. A giant G5 III star has radius of $10 R_\odot$ and mass $3 M_\odot$. A B5 I supergiant has radius $30 R_\odot$ and mass $25 M_\odot$.

6.3 The dynamo model and stars with very low L_x

The magnetic field of the Sun is thought to be generated by a dynamo driven by differential rotation at the diffusive-convection zone boundary. Convection carries the field to the surface, where it erupts in closed loop structures in which the coronal plasma is heated and confined. Surface magnetic activity appears necessary for a hot corona. As most main sequence stars age, rotation is slowed by emission of a magnetised coronal wind, and the star swells to become a red giant. These giant stars should rotate slowly, if at all, and coronal emission should disappear. Slow rotation might suppress field generation, or if a field is produced and should rise to the surface, the low gravity and/or the strong cool wind might prevent the heating and confinement of plasma at the surface.

No coronal signs, optical or X-ray, had been detected from red giants cooler than class K1, so this part of the H-R diagram was proposed as a cut-off for coronal X-rays. This is the *Linsky-Haisch dividing line*, and there have been a number of attempts to find emission from the ‘coronal graveyard’ on the cool side of this line. ROSAT and Chandra pointings (e.g. Ayres *et al.*, 2003) have yielded

the detections and limits listed in Table 6.4. It is interesting that for Arcturus and Aldebaran, $L_x/L_{\text{bol}} \leq 5 \times 10^{-11}$, compared with $\sim 10^{-7}$ for the Sun. This absent or very low emission shows that either the coronal temperature is $\leq 10^5$ K or that the corona is buried under a cool, dense, wind-forming region.

These are difficult observations. The stars are optically very bright and the X-ray detectors used have some sensitivity to IR/optical/UV radiation. Optical blocking filters must be thin to transmit soft X-rays, and some light gets through. The Chandra ACIS detector has a ‘red leak’ and the Chandra HRC a ‘blue leak’, where a minute fraction of optical photons are registered as X-rays. Pre-flight calibration exists but is difficult to apply accurately to actual observations. Thus some detections are given as upper limits. Another difficulty concerns possible unseen binary companions. A dwarf M star close to a giant can be undetected optically yet a clear source of X-rays. Fortunately, the Chandra spatial resolution is good enough so that observations of nearby stars should be free of this problem.

The main sequence (dwarf) stars also are weak X-ray emitters in two regions of the H-R diagram. There is no convection in stars earlier than type A, so going from faint to bright stars, there is

Table 6.4 Stars with very little coronal X-ray emission.

Star	Name	V (magnitude)	M_V (magnitude)	Spectral type	L_x (ergs s $^{-1}$)
α Lyr	Vega	0.0	0.5	A0 V	$< 1.1 \times 10^{25}$
α Boo	Arcturus	-0.1	-0.2	K2 III	1.5×10^{25}
α Tau	Aldebaran	0.8	-0.7	K5 III	$< 7 \times 10^{25}$

a gap in the mid-A range where the convective-driven dynamo stops and the production mechanism for early stars is weak. Thus Vega (A0 V), if it emits X-rays at all, is an extremely faint source. The second region is at the faint end of the main sequence. As stars become less massive, X-ray emission becomes very weak for stars cooler than type M5.

boxes that were themselves of a similar type. The HEAO-1 source H0324+28 was noted to be near the well-known and bright optical variable UX Ari (Walter *et al.*, 1978). It was in this way that these strong soft X-ray sources were identified with the class of stars known as the RS CVn systems, of which UX Ari is an important member. The Einstein survey soon confirmed these identifications and found X-rays from more distant members of this group.

6.4 Coronal emission from binary systems

Although emitting X-rays by the same mechanism, some stars are considerably brighter than the Sun. The first of these was discovered in 1977, before the Einstein survey, by the soft X-ray all-sky survey of HEAO-1. (Uhuru had surveyed the X-ray sky in 1971–1972, but coronal thermal radiation was too soft to register in the detectors. Low-energy sensitivity was crucial for studies of stellar coronae which have temperatures of a million to a few million degrees. At these temperatures most of the X-rays emitted are below 1 keV.

Some HEAO-1 detectors were equipped with windows as thin as 1 μm so X-rays could be detected in the range 0.2–2 keV. Within a couple months after launch, several new, strong soft X-ray sources had been discovered. The spectra of these sources indicated that the X-rays originated in a hot gas at about 10 million degrees. It was possible to determine positions of bright sources to about 0.1° which was not very good when it comes to hunting for an optical counterpart (e.g. look back at Fig. 1.4, which shows the star field around Sco X-1!). However, with several X-ray sources that seemed to have a common origin or mechanism, it was possible to search for stars in the error

Box 6.3 Variable star names

The naming system for variable stars is a classic example of failure to anticipate future needs. The bright variables, such as Mira, already had names. Newly discovered variables were given letter designations (starting with R) followed by the possessive name of the constellation. Thus new variables in Andromeda became R Andromedae, S Andromedae and so on to Z Andromedae. Needing more names, double letters were used: RR, RS, ... RZ, SS, ST, ... ZZ. Needing still more, AA, ..., AZ, BB, BC, ... QZ (but with J omitted) were called into service. Finally, appreciating an open-ended situation, the next variables were called V335, V336 and so on.

This designates individual objects. If a number of objects are found to have certain similar features (such as characteristic colours or timescales of variability), then that group of objects is named after the first one that was found with that particular characteristic. In the case we consider the star RS CVn was just such a prototype. CVn is short for the constellation Canes Venatici, and RS indicates that it is a variable star which is listed in the massive Russian General Catalogue of Variable Stars (or GCVS).

Table 6.5 Characteristics of the RS CVn binary systems.

Binary period	1–14 days
Spectral type	K0 IV/V + main sequence star
Spectral features	Ca, H emission lines
Masses	1–1.5 M_{\odot}
Optical brightness	sixth to ninth magnitude
Optical wave amplitude	0.1–0.3 magnitude
Wave migration period	years
Colour	UV and IR excess
Variability	radio and X-ray flares

6.4.1 RS CVn systems

The characteristics of the RS CVn systems were laid down in a now classic paper (Hall, 1972). Their main properties are summarised in Table 6.5, and a list of the bright X-ray-emitting RS CVns is compiled in Table 6.6. Figure 6.10 shows a schematic of RS CVn itself.

6.4.2 Giant starspots on RS CVn stars

At first sight these RS CVn systems seem to have little in common with our Sun and its coronal X-ray emission. RS CVns are close binary systems, one of whose components is ageing and slightly evolved. The clue to the nature of their activity comes from the variability. Figure 6.10 shows the

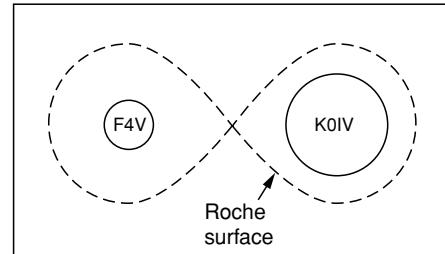


Fig 6.10 Scale drawing of the RS CVn binary system. Although a detached binary (i.e. there is no mass transfer taking place between the stars), they are close enough to be locked rotationally. Each star rotates once per binary cycle.

size and separation of the stars, and Fig. 6.11 shows the light curve of RS CVn in 1975.

Both stars pass in front of each other at opposite sides of the orbit, producing the primary and secondary eclipses that are easily visible (the primary eclipse is almost a full magnitude in depth). The eclipses are a straightforward way of indicating a binary star but represent a selection effect in the way these stars were found in the optical. The X-ray emission of RS CVns, however, suffers no such effect and will enable many more non-eclipsing systems to be found. Given the sizes of the stars, we would expect only about 10 per cent of such systems to be eclipsing as observed. This implies that there are about 10 times more objects of comparable optical brightness in this class to be found.

Table 6.6 Properties of X-ray bright RS CVn systems and other active binaries.

Name	Distance (pc)	Visual magnitude	Rotation period (days)	Rotation Spectral type	L_x (0.2–3 keV) (10^{30} ergs s $^{-1}$)
UX Ari	50	6.5	6.4	K0 IV + G5 V	21
HR 1099	36	5.7	2.8	K0 V + G5 V	26
RS CVn	150	8.4	4.8	K0 IV + F4 III	19
AR Lac	40	6.9	2.0	K0 IV + G2 IV	15
LX Per	145	8.1	8.0	K0 IV + G0 V	6
HK Lac	139	6.5	24.4	K0 III + F1 IV?	14
σ Gem	37	4.3	19.5	K1 III + G V?	8
Capella	13	0.1	104.0	G8 III + G1 III	4
Algol	27	2.1–3.3	2.9	B8 V + K2 IV	5

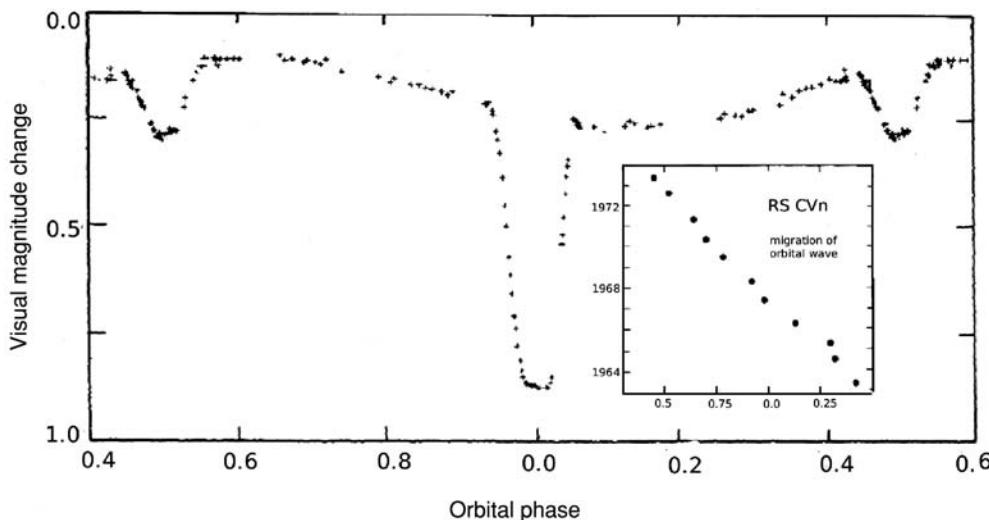


Fig 6.11 Optical light curve of RS CVn. Both primary and secondary eclipses are easily visible, but note the gentle undulation or wave in the light curve. At the time of this observation in 1975, the wave peaked at orbital phase ≈ 0.6 , but the wave drifts or migrates and takes about 10 years to return to the same binary phase. This is comparable to the Sun's own activity cycle.

The slight undulation or wave in the light curve is very important and is visible in Fig. 6.11. This wave, surprisingly, does not have exactly the same period as the binary period. If the position of the minimum in the wave is plotted as a function of binary phase, this changes smoothly with time. It takes about 10 years in RS CVn for this migration to march completely through one cycle and back to the phase at which it started. Whilst this wave migration is taking place, the amplitude of the wave is varying slowly with time. In RS CVn the wave changes from an amplitude of 0.20 to 0.05 magnitude on a timescale of 20–25 years. This timescale is considered to be analogous to the solar 11-year sunspot cycle and the wave to be caused by the presence of giant starspots on the surface of one of the two stars in the system, as imagined in Fig. 6.12.

These starspots have to be gigantic to explain the amplitude of the wave that is observed. They must cover a significant fraction of the stellar surface. For comparison, sunspots occupy only about 0.1 per cent of the Sun's surface at solar maximum, whereas these starspots must cover typically 20 per cent of the star's surface. More importantly, they have to cover 40 per cent of one hemisphere since the starspots cannot be uniformly distributed but must be concentrated

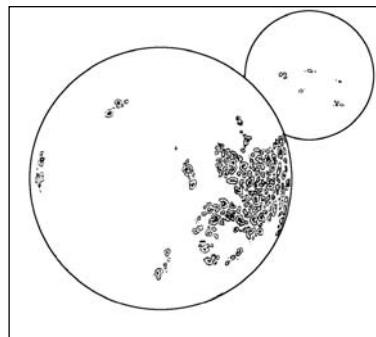


Fig 6.12 Artist's conception of giant starspots and the RS CVn binary. These are all scaled-up versions of actual features seen on the Sun (see Fig. 6.7).

on one-half of the star to explain the light curve. And just like sunspots, which have cycles of activity that change every year, so do the starspots. The wave of the RS CVn system λ Andromedae changes its overall shape and structure on timescales of 1 year.

Observations of the wave and the binary motion show that these starspots are associated with the subgiant member of the RS CVn binaries (which is a K0 IV star in almost every system). This star has started to evolve off the main sequence because it is running out of its nuclear fuel (hydrogen) at the centre of the star.

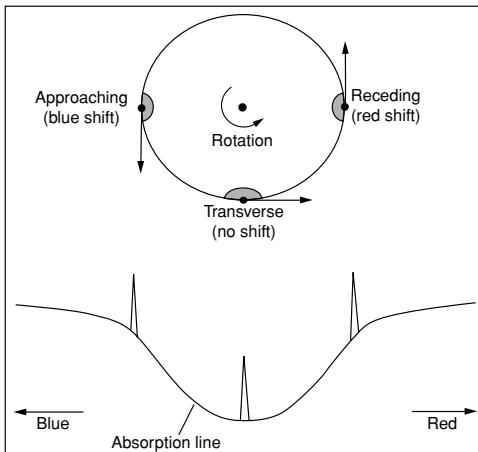


Fig 6.13 Doppler imaging of starspots. This shows a rotating star with a single large spot at three different times. The activity associated with the spot produces an emission line, whereas the rest of the stellar surface produces only absorption (broadened because it is coming from all parts of the surface which are travelling at different velocities relative to our line of sight). The emission line appears in the centre of the broad line when the spot is travelling directly across our line of sight and at Doppler-shifted positions when the spot is travelling towards or away from us, at the edges of the stellar disc. By studying the line profile at different times, the position of the spot on the stellar surface can be determined, hence the term *Doppler imaging*.

It starts to expand (on its way to becoming a red giant) and develops a much more substantial and deeper convective envelope which generates the turbulence.

The starspots can explain both optical and X-ray characteristics of the RS CVn systems. As footprints for even larger active regions, they account for the greatly enhanced chromospheric activity (calcium and hydrogen emission lines) and for the bright X-ray coronae. The RS CVn coronae are $\sim 10^4$ times brighter than that of the Sun.

6.4.3 Doppler imaging of starspots

Given that it is not possible to resolve the stellar discs, the starspots cannot be seen. A technique called Doppler imaging, however, has enabled the determination of the distribution of starspots on a stellar surface. The technique is based on the Doppler effect and is explained in Fig. 6.13.

The starspot positions in Fig. 6.14 are derived directly from the line profiles that are shown

under each map. The narrowness of the emission features shows that the emission regions are localised with respect to the stellar disc. Because the lines are from enhanced chromospheric emission immediately above the spots, the lines locate the spots. These lines can be resolved by high-resolution spectroscopy as part of the much wider absorption line profiles that come from the entire stellar disc.

Giant starspots will presumably support correspondingly giant coronal loops (scaled from those seen on the Sun, as in Fig. 6.4). This large active region has a hot plasma at about 10 million degrees contained by the magnetic field. It is possible to scale very roughly the theoretical work done on the solar coronal loops up to the sizes required for the RS CVn systems. The temperature of the RS CVns is known quite accurately from X-ray spectra. The emission lines are only explainable by a very hot gas. For Capella, there is emission by a range of Fe ions, indicating the presence of at least two spectral components, with temperatures of a few million and 10–20 million degrees or greater.

6.4.4 Eclipse imaging of coronal structure

If the wave in the light curve of RS CVn systems is attributable to giant starspots on one hemisphere of the active star, then we expect the X-ray-emitting regions to be similarly distributed. This can be tested by measuring the X-ray emission of an X-ray bright, eclipsing RS CVn system in which the presumably less active component eclipses the active star. By observing the X-ray output during the eclipse, it is possible to determine the structure of the X-ray-emitting regions.

A good system for such a study is AR Lac, although its two components are both subgiants, suggesting that both may be X-ray active to some extent. This RS CVn system has a 2-day period and the properties summarised in Table 6.6. In July 1984 and June 1993, respectively, EXOSAT and ASCA observed AR Lac, both for a continuous period of 2 days, essentially covering one complete orbital cycle. The X-ray spectrum requires at least two temperatures. The coronal emission has a low temperature component with temperature $\approx 7 \times 10^6$ K and a high-temperature component at $\approx 2 \times 10^7$ K. The different energy ranges of the

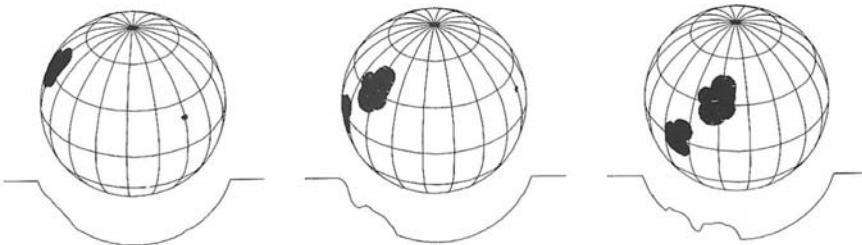


Fig 6.14 Doppler imaging of the starspots on UX Ari. The rapidly rotating star produces a normally broadened absorption line (shown schematically here as a dishlike feature). However, the starspots are active regions confined to particular areas of the star's surface and give rise to emission at a single Doppler-shifted velocity rather than a range of velocities. The emission from each starspot starts to fill in part of the absorption line, giving rise to the structure in the line profile. By following this structure as a function of time, it is possible to map the positions of spots on the surface and to track their movement across the star.

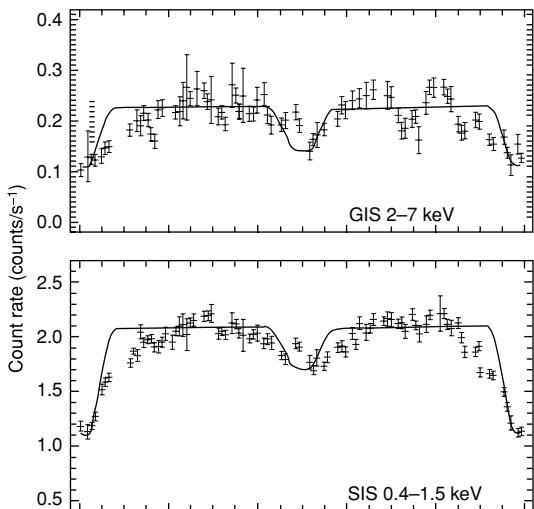


Fig 6.15 The ASCA-measured X-ray light curve of AR Lac for one orbital cycle. The GIS detector sees only the higher-temperature regions, and the SIS detector sees low-temperature emission also. A few brief flares have been removed. The solid curve is the optical light curve for comparison. The times of primary and secondary eclipse correspond to phases 0° and 180° , and both are approximately 8 hours in duration (from Siarkowski et al., 1996).

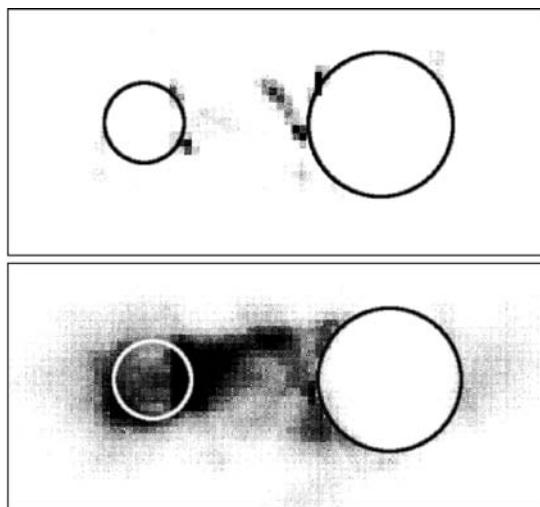


Fig 6.16 A map of the AR Lac coronal emission derived from the X-ray light curve (in Fig 6.15). The smaller star, at the left, is the G2 IV primary and the larger is the K0 IV secondary. The deeper primary eclipse occurs when the K star occults the G star. (top) High-temperature regions, which are localised. (bottom) Low-temperature regions, which are more diffuse and strongest in the space between the two stars. Dark and light circles showing positions of the stars do not imply emissions from the surfaces (from Siarkowski et al., 1996).

ASCA detectors make the GIS sensitive to essentially only the high temperature coronal component and the SIS sensitive to both high and low components equally.

The ASCA X-ray light curve is shown in Fig. 6.15. The spatial extent of the coronae, derived from the light curves, is shown in Fig. 6.16. We note that this map is not unique. Other distributions

are possible, and the extended structure between stars is not always required. Nevertheless, this was the most probable solution. The technique used was the *maximum entropy method*, which derives the smoothest distribution of X-ray brightness across the stars consistent with the observed light curves.

Because the X-ray flux is not zero at the minimum of the primary eclipse, there must be

an extended component larger than the size of each star. It could be a halo surrounding both stars or only associated with the poles of the larger K star. The strong primary eclipse requires extended emission from the facing hemispheres of both stars. The extent of this emission is uncertain, but some emission from compact regions not far from the surface of both stars is required. As evident from the high-temperature GIS map, the corona of the K star is hotter than that of the G star.

The bright, compact regions are likely to be coronal loops similar to those on the Sun, but at a higher temperature and over 1000 times larger than those on the Sun. It is now becoming clear that the magnetic fields play a crucial role in the activity of the RS CVn systems, and this strongly supports the magnetic heating model. Why these loops are so much larger and more powerful than in the Sun, and what energy source is driving them, are open questions.

6.5 High-resolution X-ray spectra

The first observations of stellar coronae were marginal detections, and spectra were fit with single-temperature models. Later observations with more sensitive low-resolution detectors proved these models inadequate, and two-temperature fits were applied to the data. Most RS CVn stars required both a high- and low-temperature component. At this time, everyone realised that, since the observations integrate over many coronal structures, the emission would be characterised by a continuous range of temperature. Modern data are good enough to measure this.

In 1999 Chandra and XMM carried sensitive high-resolution spectrometers into orbit and started a new era of coronal diagnostics. This section shows a few observations and results.

Figure 6.17 shows spectra of five stars measured by the Chandra LETG. The most active star is at the top. Temperatures decrease going from top to bottom. This is apparent from the relative intensity of various emission lines. Ne X is strong for UX

Ari and C IV for α Cen A&B. The O VIII/O VII ratio decreases from top to bottom. The most luminous stars have the highest temperatures. Luminosities range from 1.5×10^{31} ergs s⁻¹ for UX Ari down to 3×10^{27} for α Cen A&B, about that of the Sun at solar maximum. The intermediate stars ϵ Eri (K2 V) and AD Leo (dM4.5 Ve – a flare star) have $L_x \sim 4 \times 10^{28}$ ergs s⁻¹, 10 times that of the Sun.

Capella is an active, X-ray bright binary with a 104-day period. We list it in Table 6.6 because it is close to the RS CVn stars in nature. It was the first non-solar coronal source observed in X-rays, discovered with a rocket-borne detector in 1974 (Chapter 1), and observed by many missions since. Colour Plate 12 shows a sum of several Chandra observations. This shows the image produced by the Chandra HETG spectrometer. Colour Plate 13 shows part of the extracted spectrum. Individual lines are well resolved and indicate elemental composition, temperature and density of the emitting plasma. Let us first consider temperature.

The observation of a particular ion species implies a particular temperature range (as shown in Fig. 2.2). Additionally, the ratios of H-like to He-like ions for a single element further define a temperature. For Capella the ratio of O VIII/O VII requires $T = 3.8 \pm 0.2 \times 10^6$ K, and the ratio Si XIV/Si XIII requires $T = 9.5 \pm 0.5 \times 10^6$ K. The corona is not isothermal (and was not expected to be). This information can be used to construct a differential emission measure (DEM) for the star. Recall that the *emission measure* (EM) is a function of temperature, and for a volume V of isothermal plasma of density n ions cm⁻³, $EM(T) = n_e n_i V$. On the other hand, the $DEM = n_e n_i \Delta V / \Delta T$ and shows the relative contribution of coronal plasma at different temperatures.

Lines from He-like ions are triplets, as shown in Figs. 6.18 and Colour Plate 13. The first excited state of the He-like atom has three levels all at slightly different energies, dependent on the angular momentum of the excited electron. These three states are populated by photon absorption, and they decay by emitting photons at three slightly different wavelengths (around ≈ 22 Å for O VII). These transitions are, in order of increasing wavelength, called *resonance*, *intercombination* and *forbidden*. These lines are easily detected and

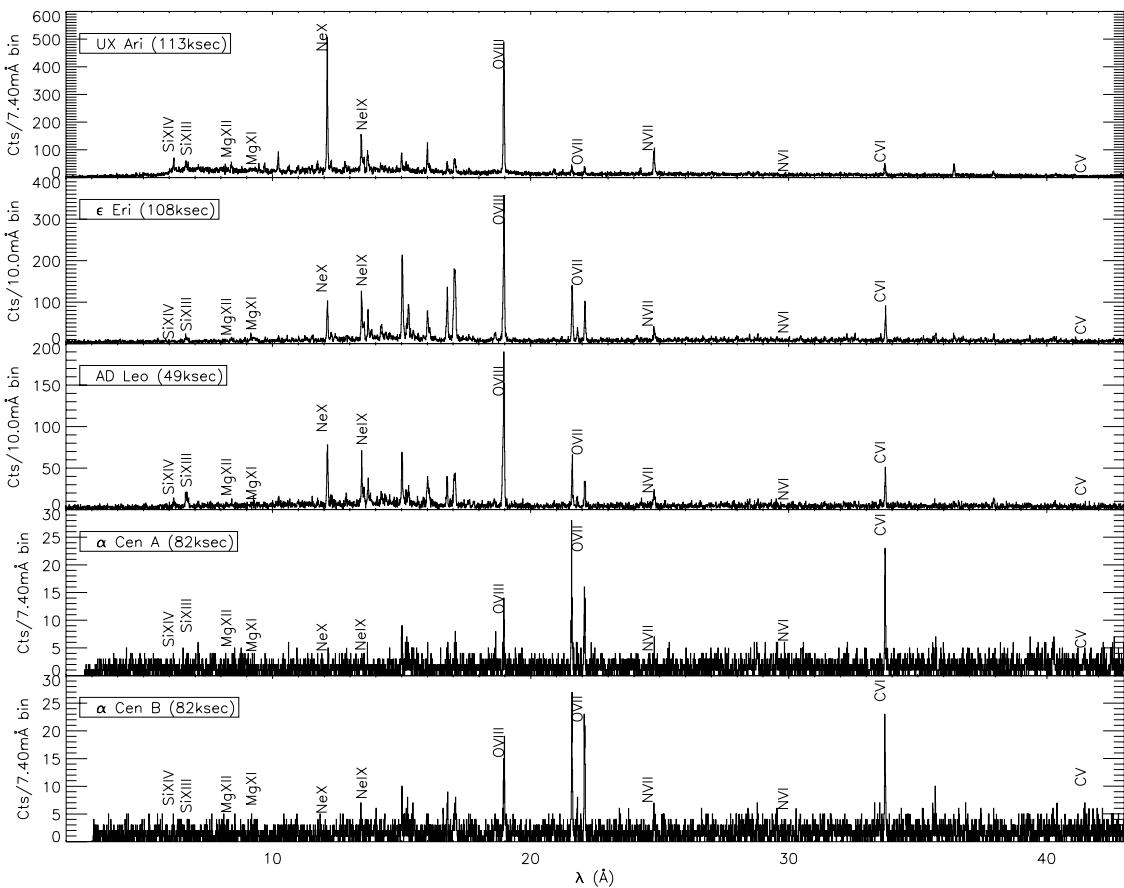


Fig 6.17 Chandra LETG spectra of five stars. The most active is the RS CVn star at the top; least active are α Cen A and B at the bottom. Because this is a diffraction grating, the x axis is in wavelengths. The more luminous stars are hotter, and shorter wavelength lines are stronger. Emission lines are labelled. Note the O VII lines used to determine coronal density (from Ness et al., 2002).

resolved by Chandra and XMM spectrometers. Sometimes, however, there is blending with lines from other elements. The O VII triplet is usually clear of contamination and is useful in density diagnostics. The forbidden line comes from a long-lived excited state. If the electron density is high enough, collisions will cause a transition from the forbidden state to the intercombination state, which then decays by emitting a photon. As collisions become more frequent, the forbidden line becomes weaker and the intercombination line stronger.

Thus the ratio of the intensity of the two lower-energy lines is an indication of density but only over a limited range. O VII can be used in the range

10^{10} to 10^{11} electrons cm^{-3} and Mg XI in the range 10^{12} – 10^{14} . Figure 6.18 shows a detection of the O VII triplet from Capella. At the low-density limit, the ratio f/i is 3.95. The ratio from the data in Fig. 6.18 was $f/i = 2.9 \pm 0.4$, implying a density of 0.8 – 2×10^{10} electrons cm^{-3} . The density derived from Mg XI triplet ratios (Colour Plate 13) gave an upper limit of 7×10^{11} electrons cm^{-3} .

The DEM for Capella is shown in Fig. 6.19. This and the density information define the corona. Knowing the luminosity and density, a volume can be calculated. Results show that even though the Capella observation encompasses two stars, the derived coronal loop sizes are small compared to the stellar radii.

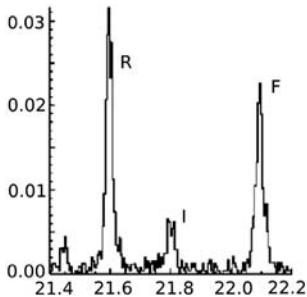


Fig 6.18 The He-like O VII triplet from a Chandra HETG observation of Capella. Count rate is plotted as a function of wavelength in Å. The three strongest lines from left to right are the recombination, intercombination and forbidden lines (from Canizares *et al.*, 2000).

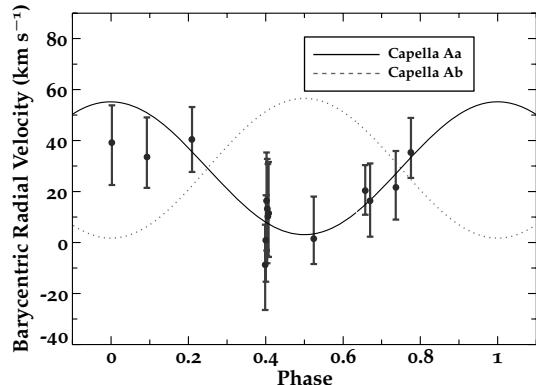


Fig 6.20 Observed Doppler velocities of emission lines from Capella as a function of binary orbital phase. They follow the motion of the G8 III primary, shown by a solid curve. The secondary star velocity is indicated by the dotted curve (from Ishibashi *et al.*, 2006).

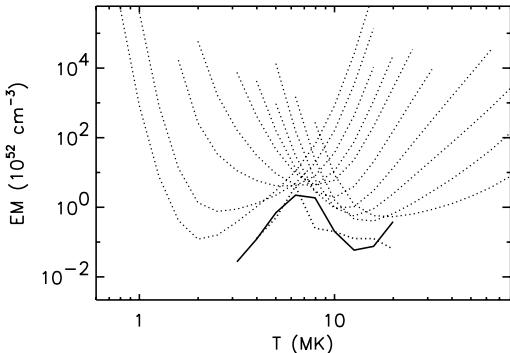


Fig 6.19 The DEM for Capella, showing that most coronal emission is from material at temperature of $5\text{--}9 \times 10^6$ K. The solid curve is derived from Fe emission lines ranging from Fe XV to Fe XXIV. Upper dotted curves show individual line emission measures. The lowest dashed curve shows an earlier DEM obtained from ASCA and EUVE observations, a measurement not far from the Chandra result (from Ness *et al.*, 2003).

6.6 | X-ray Doppler imaging

The first example is simple in concept but required care in data handling. Capella is a binary system, a G8 III and a G1 III star in a 104-day orbit. There are no eclipses. The orbital velocity (projected on the line of sight) of the two almost equally massive stars is $\approx 30 \text{ km s}^{-1}$. Doppler shifts of the emission lines can be used to determine which star they come from. It is very important, however, to first correct to the Earth-Sun barycentre to remove the

$\approx 30 \text{ km s}^{-1}$ shift owing to the orbital velocity of the Earth. Figure 6.20 shows the result from 14 Chandra HETG observations spread over a 6-year interval. The derived Doppler motions fall on the velocity curve of the primary star (G8 III), showing that this is the dominant source of coronal emission.

The second example is the contact binary 44i Bootis of spectral type G0 V. The orbital period is 6.4 hours and orbital velocities of the two components are 100 and 200 km s^{-1} . Inclination is 72° , and there are two partial eclipses each orbit. This system was observed for two complete cycles with the Chandra HETG (Brickhouse *et al.*, 2001). The line width was narrow enough to preclude any diffuse extended emission, and so the corona is localised. The line centroids were found to shift with orbital phase and to track the motion of the primary star (exactly as shown in Fig. 6.20 for Capella). The X-ray light curve showed only one eclipse per cycle, so the emission is not equatorial. The observed modulation was also small, implying that most of the corona was always visible. The conclusion was that most of the coronal emission originates in the polar region of the primary star. Because active regions on the Sun are always equatorial, this was a surprise. Colour Plate 11 shows an artist's impression of this system.

A third example is the rapidly rotating single star FK Com, a type G III star with radius $\approx 9 R_\odot$,

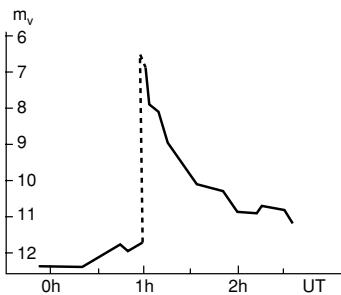


Fig 6.21 An optical flare of UV Ceti observed 25 September 1952. The star was 6 magnitudes brighter at the peak of the flare, a factor of ~ 100 increase in luminosity (from Petit, 1954).

spin period of 2.6 days, and equatorial velocity $v\sin i = 160 \text{ km s}^{-1}$. Optical Doppler imaging indicated active regions at high northern and low southern latitudes with concentration over $\sim 180^\circ$ of longitude. A Chandra HETG observation measured a Doppler shift in the X-ray emission lines, showing that the X-rays originated from the same hemisphere. The width of the lines was wider than anticipated. Because line width depends on rotational velocity, which increases with height in the stellar atmosphere, this result implies that the X-rays originate high above ($\sim 0.15 R_*$) the surface. Thus both optical and X-ray Doppler results can determine gross coronal structure.

6.7 Flare stars

A stellar flare is a sudden, often dramatic increase in brightness lasting usually minutes but sometimes hours or days. The first flare stars discovered were all dMe stars. The best known is probably UV Ceti, which is the prototype for this class of object. UV Ceti is the cooler member of a nearby binary system (M5.5 V/M6.0 V, period 26 years, distance 2.7 pc) and at 13 mag is visible only through a telescope. The optical luminosity of flares on UV Ceti often exceeds that of the star, a marked contrast with a flare on the Sun. However, because UV Ceti is so dim, a solar-type flare is much easier to see. Figure 6.21 shows one of the earliest observations of a large UV Ceti flare.

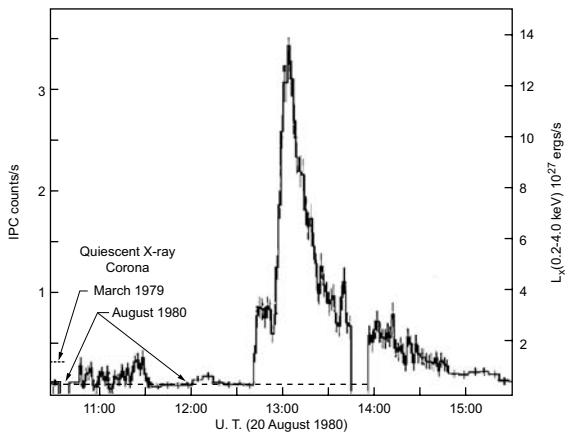


Fig 6.22 A large X-ray flare on Prox Cen observed by Einstein in the energy range 0.2–4 keV. Emission at maximum was $1.4 \times 10^{28} \text{ ergs s}^{-1}$. A mark on the left side notes the quiescent emission from a previous observation. This varied between 1.3 and $0.3 \times 10^{27} \text{ ergs s}^{-1}$ (from Haisch et al., 1983).

It is remarkable that flare stars are otherwise some of the smallest and dimmest objects known. Spectral type M corresponds to an intrinsic optical brightness at least 100 times fainter than the Sun and masses of only a fraction of that of the Sun (hence the *d* for ‘dwarf’). M-type stars are, however, the most common object in our Galaxy, probably outnumbering all other stars put together. Indeed, 40 of the nearest 63 stars are M type. The *e* suffix is added because emission lines have been seen in their optical spectra, the first indication of some kind of activity.

Observing optical flares requires patience and, usually, short integration times since they are often over within a matter of minutes. UV Ceti was not seen flaring until 1948, although archival plates at Harvard were searched to show that similar flares had in fact been occurring since 1900. There are now more than 100 flare stars known.

Like solar flares, stellar flares emit copious X-rays. Given the similarity of stellar flares to solar flares (but often scaled up!), several satellites took part in early campaigns to search for X-rays during flares. The Dutch/U.S. satellite ANS was successful, and X-rays were found from two optical flare stars (even before the discovery of X-radiation from the RS CVn systems). Before the Einstein Observatory,

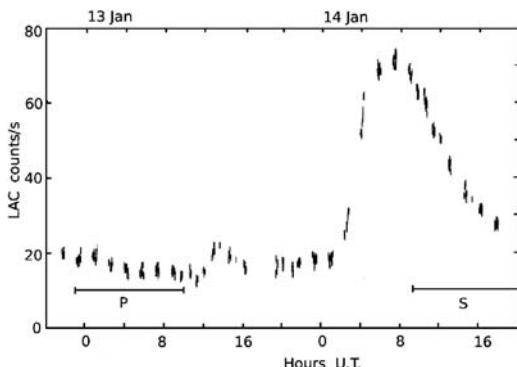


Fig 6.23 Two X-ray flares on Algol observed in 1989 by Ginga in the range 1–18 keV. The light curve is broken because there was an Earth occultation of the source every 100 min. Horizontal lines mark the primary and secondary eclipses of Algol. There was a small flare at ≈ 1230 UT on 13 January, and the large flare at ≈ 0200 UT on 14 January is obvious. Emission at the peak of the large flare was $\approx 1 \times 10^{31}$ ergs s^{-1} . Preflare quiescent emission was 3×10^{30} ergs s^{-1} (from Stern et al., 1992).

none of the known flare stars were detected as continuous X-ray sources. Einstein observed several flare stars. Figure 6.22 shows an observation of Proxima Centauri during which a large flare occurred in the middle of the session. It was also noted that flare star quiescent emission was quite variable, a sign of continuous coronal activity. This is indicated in the figure.

During the 1980s large X-ray flares were also discovered from other stars quite different from the dMe stars. EXOSAT and Ginga, for example, measured large flares from β Per (Algol), a famous naked-eye B8 V/K2 IV eclipsing binary with a 2.86-day period. Durations of the large flares were hours. These detectors were not efficient at X-ray energies below 1 keV, but there was little loss of sensitivity. The flare spectra were hard, and the Fe emission line at 6.7 keV was strong.

The large flare from Prox Cen in Fig. 6.22 was 60 times more luminous than the class M8 solar flare shown in Fig. 6.5, or about 3 times more luminous than really large solar flares (class X20). The large Algol flare was ≈ 2000 times more luminous than a solar X20 flare. The Algol K2 IV star is the source of the coronal X-rays (because B8

stars do not have active coronae). Because the Algol flare in Fig. 6.23 was not affected by the secondary eclipse (when the B8 V star passes in front of the larger K2 IV star), it was probably a large structure. Indeed, loop models require a size comparable to the radius of the K2 subgiant ($\sim 3 R_{\odot}$).

Figure 6.24 shows a modern XMM observation of a large X-ray flare on Prox Cen. Peak X-ray luminosity was 4×10^{28} ergs s^{-1} (3 times larger than the Einstein-observed flare in Fig. 6.22) and the XMM RGS simultaneously recorded spectra. The figure shows the Ne IX and O VII triplets used for density analysis. Differences in the f/i line ratios can be easily seen in this illustration. The O lines are well resolved; the Ne lines are not. Three-line fits are indicated by the histograms. The X-ray derivative in Fig. 6.24b is there to illustrate the Neupert effect, a relation between the rise of X-ray emission and radiation at other wavelengths. The U-band light (and sometimes radio emission) is produced by an initial burst of high-energy electrons which heat coronal plasma. This heated material expands into the loop structure and radiates X-rays. The rate of increase in X-ray emission follows the rate of energy deposition signaled by the U-band light.

There is every indication that these large stellar flares are analogous to solar flares. They are a product of magnetic activity caused by stellar rotation and convection. Type dM stars have deep convection zones, and close binaries like Algol rotate rapidly. This picture is reinforced by the observation of large flares from pre-main sequence stars, which are the subject of the next section.

Algol is a binary with strong coronal emission similar to the RS CVn stars where flares are also common. The key feature is rapid rotation. In close binary systems, this is easy to produce because each star rotates synchronously with the orbital period (in exactly the same way that the Moon rotates once in each lunar month so that it always keeps the same face towards the Earth). The orbital periods are so short (days) that the rotational speed is correspondingly fast when compared, say, with the rotation period of the Sun, which is about 27 days.

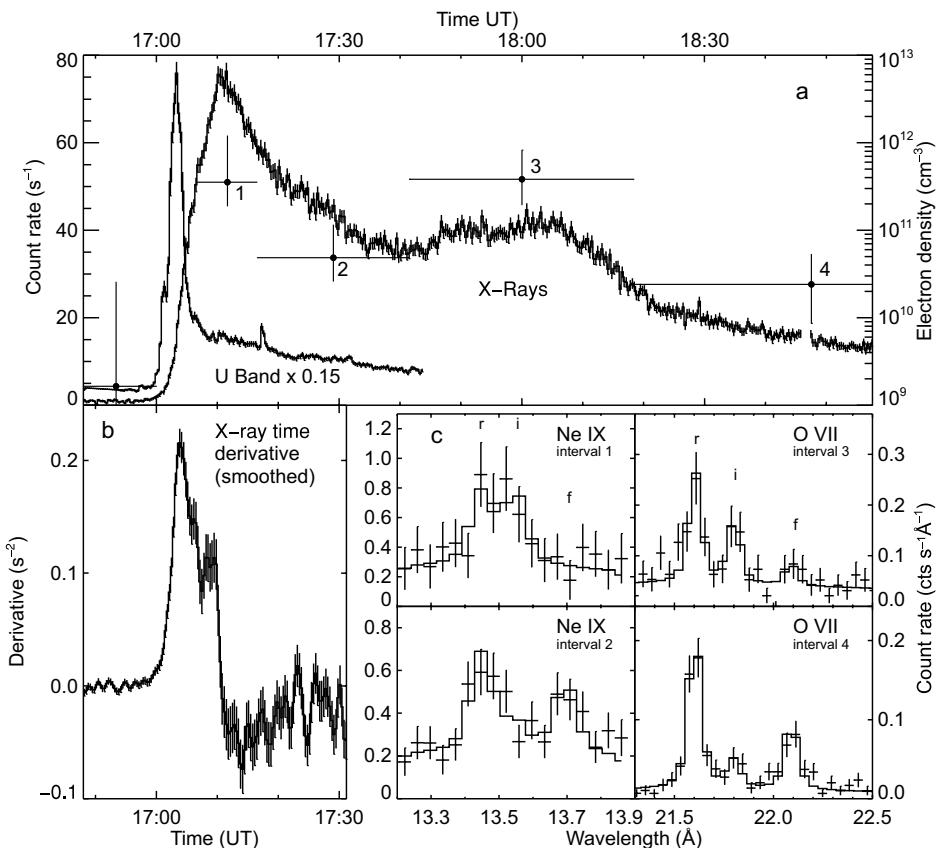


Fig 6.24 A large flare on Prox Cen observed by XMM on 12 August 2001. The EPIC PN detector recorded X-rays in the 0.15–10 keV band, the RGS spectrometer observed the dispersed spectrum and the optical monitor measured the U-band optical flux – all simultaneously. This is a nice demonstration of XMM capability. (a) Light curves over a 2-hour period. (b) Comparison with the U-band profile. (c) Two He-like triplets used to determine plasma density, shown on the right-hand scale (from Güdel *et al.*, 2002).

6.8 | Young stars

We now consider stars at an early stage of their life. Stars form from large interstellar clouds of gas and dust that collapse under their own gravity. As part of the cloud collapses to form a star, it is first heated by the release of gravitational energy. At a later time the star becomes dense enough for thermonuclear reactions in the core, and it settles onto the main sequence to start its life as a ‘normal’ star. To reach this point, the star must shed angular momentum. A disc forms around the protostar. The protostar accretes and sheds material, and at this point we have a *pre-main*

sequence or PMS star. Magnetic fields couple the rapidly rotating PMS star with its surroundings. PMS stars are strong, active X-ray sources.

Theoretical calculations of such a collapse are very difficult, but observationally, we are seeing them earlier and earlier in the life cycles of stars. This has particularly come about with the advent of sensitive infrared (IR) telescopes and instrumentation which can detect the very cool clouds and the protostars embedded within them. The X-rays, too, can penetrate obscuring dust, enabling to see blanketed, very young stars. The collapse of a large cloud forms many stars, all at about the same time and in the same place. The result is a cluster of stars such as the Hyades or the Pleiades.

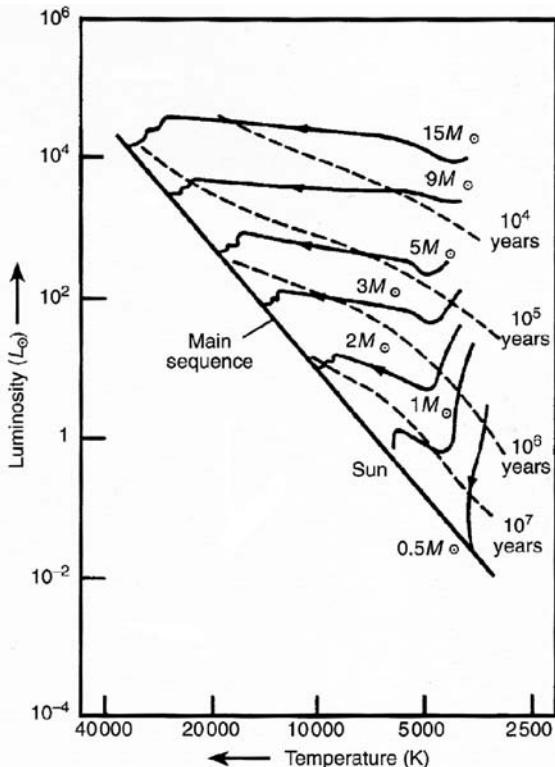


Fig 6.25 Tracks of young evolving stars in the HR diagram. Solid curves show tracks followed by PMS stars of different masses. Dashed curves show the time required after the star is formed to reach a particular point. Thus, in 10^7 years, stars more massive than $2 M_\odot$ have reached the main sequence; stars less massive have not.

Einstein and ROSAT surveyed many star-forming regions and young clusters, and hundreds of PMS stars were found to emit X-rays with $L_x = 10^{29}-10^{31}$ ergs s $^{-1}$. The dependence of luminosity on age shown in Fig. 6.28 was established. Variability of most PMS stars was strong, and large stellar flares were common.

6.8.1 The Chandra Orion survey

One of the most well known and most spectacular regions of star formation is in the Orion Nebula. This was the target of a 838-ks (≈ 10 -day) survey by Chandra with the result shown in Colour Plate 15. This region, with heavy absorption and a high density of stars, was well suited for Chandra observations. Sixteen hundred X-ray pointlike sources were found. The exposure was deep enough to

detect the faintest T Tauri stars, and with a few exceptions, all the stars known to be PMS stars from optical observations were found to emit X-rays. A big advantage of this survey is that all

Box 6.4 Evolution of pre-main sequence stars

The most massive stars condense so quickly that they are never seen as PMS stars. When they appear, they are on the main sequence as O stars. Less massive stars spend 10^5-10^7 years as PMS stars powered by gravitational contraction. They follow the tracks shown in Fig. 6.25 to the main sequence, where H burning in the core supplies the power. A G star, like the Sun, starts out cooler but much brighter than the Sun. When younger than 10^7 years, it appears as a K giant.

PMS stars with mass $\leq 2 M_\odot$ are called *T Tauri stars*. They have spectral types F to M. PMS stars with mass $2-8 M_\odot$ are *Herbig Ae/Be stars*.

stars are at the same distance, so relative luminosities are well determined. Varying absorption across the region, however, must be taken into account.

Figure 6.26 shows the X-ray luminosity of the Orion T Tauri stars compared to an optical sample of main-sequence stars. The T Tauri stars are all X-ray brighter, and for any one type, there is a factor of 100 variation in the observed X-ray flux. It was also found that T Tauri star atmospheres are hotter than those of MS stars.

Although they spin rapidly, the enhanced T Tauri brightness does not arise from rotation. Figure 6.27 shows the effect of rotation on both MS and TTauri stars. The faster-rotating MS stars have higher X-ray luminosities, whereas the T Tauri stars do not show this dependence on rotation. There must be other factors, such as properties of the convection zones, that produce a high level of magnetic activity. (Note that Fig. 6.27 shows L_x/L_{bol} . For a given mass, the T Tauri stars have a higher L_{bol} than the corresponding MS stars. The two classes are not separated on the plot as they were in Fig. 6.26.)

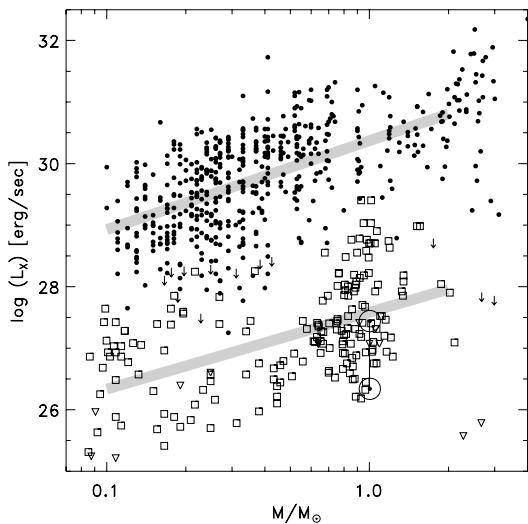


Fig 6.26 The X-ray luminosity of PMS stars in the Chandra Orion survey as a function of stellar mass. Solid points are T Tauri stars. Open squares show MS stars. Open circles mark the Sun at solar minimum and solar maximum. Grey lines are fits to the two groups (from Preibisch et al., 2005).

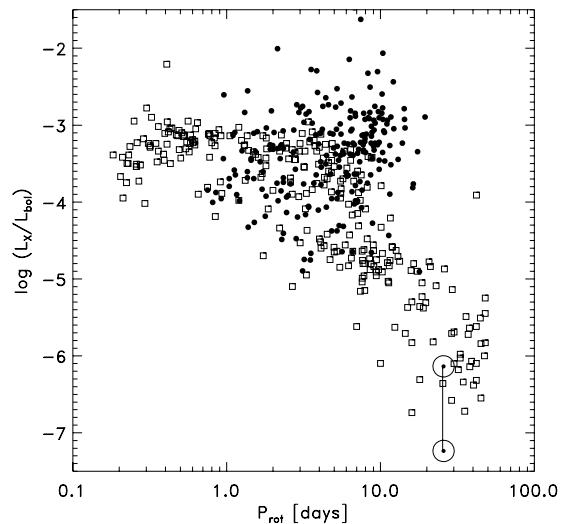


Fig 6.27 Fractional X-ray luminosity vs rotation period for PMS stars in the Chandra Orion survey. Solid points are T Tauri stars. Open squares show MS stars. Open circles mark the Sun at solar minimum and solar maximum (from Preibisch et al., 2005).

IR and emission line widths show that many T Tauri stars are surrounded by discs of material and that many of them are accreting. But accretion does not seem to account for the X-rays. The accreting stars actually show somewhat lower levels of X-ray emission than the rest. Furthermore, plasma temperatures are higher than expected from accretion shocks.

The enhanced X-ray emission of T Tauri stars is most likely because of high solarlike magnetic activity. This picture is reinforced by the observation that flares are a common phenomenon and that flares require magnetic heating and confinement. Because of photon statistics in this survey, only larger flares could be distinguished. Many were observed with luminosities that exceeded 10^{32} ergs s $^{-1}$, 10 times that of the Algol flare shown in Fig. 6.23! The loops required to hold the radiating plasma for these events are big. The lengths required are greater than several R_* , leading to speculation that the fields defining the loops are anchored to or interacting with the circumstellar discs.

The Orion Nebula is young. Not only are the T Tauri stars still embedded in gas and dust, but the nebula contains several early-type stars. The

brightest of these is θ^1 Ori C, of spectral type O7. The presence of a short-lived O star indicates a young cluster. Data from this survey have been used to update the luminosity-age relationships obtained from earlier work. As shown in Fig. 6.28, the T Tauri stars of Orion are 10 times brighter than comparable stars in the Pleiades.

6.8.2 Discs, jets and planets

Discs around T Tauri stars have been imaged with millimeter interferometers and by the Hubble Space Telescope. They are very large (~ 100 AU) and very cold compared to the accretion discs around white dwarfs and neutron stars that are considered in later chapters. These discs are not only important in the star-formation process but are also the birthplace of planets. There is speculation that X-rays could affect a disc in such a way that planets might be more likely to form.

The star TW Hydrae is one of the closest T Tauri stars, is well studied and is surrounded by a dust-rich gaseous accretion disc. A Chandra high-resolution spectrum was obtained when the source luminosity was $L_x = 1 \times 10^{30}$ ergs s $^{-1}$. Results were unusual and were interpreted to mean that at least some of the X-rays were due

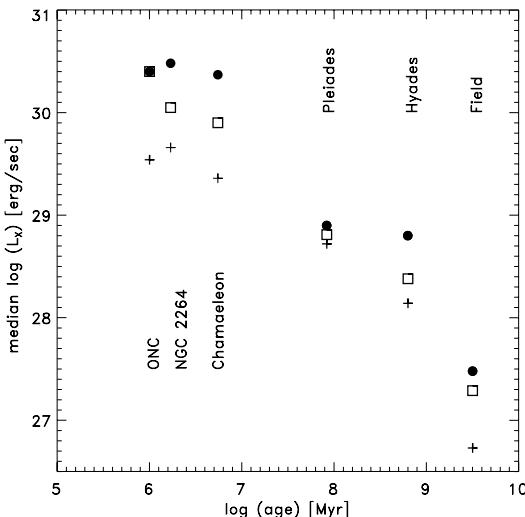


Fig 6.28 Median X-ray luminosity as a function of age for stars in young clusters. As a general rule, luminosity decreases with age. ONC is the Orion Nebula Cluster. Solid dots show $0.9\text{--}1.2 M_{\odot}$ stars. Open squares show $0.5\text{--}0.9 M_{\odot}$ stars. Crosses show $0.1\text{--}0.5 M_{\odot}$ stars (from Preibisch & Feigelson, 2005).

to accretion rather than to a magnetic corona. Colour Plate 14 shows an artist's rendition of this object and the He-like OVII triplet recorded by Chandra. The forbidden line is very weak, implying a density in excess of 10^{12} cm^{-3} (compare with figure 6.18). This is high for coronal emission and implies a rather compact source. Furthermore, the differential emission measure indicated a source with a temperature range more narrow than that of coronal emission. The model proposed is that accretion occurs in a narrow stream. Material is shocked before arriving at the star. Conditions expected in this shocked region are close to those observed. There may be X-rays from shocks in other T Tauri observations but confused by the additional presence of coronal X-rays.

T Tauri stars also have jets. The accretion disc probably collimates outflowing material along the rotation axis of the system. Because velocities of this material are $100\text{--}400 \text{ km s}^{-1}$, shocked material will be hot enough to produce X-rays. Figure 6.29 shows an X-ray observation of DG Tau which images X-rays from the central star and two jets which extend $5''$ from the star in both directions. For each jet, $L_x = 1 \times 10^{28} \text{ ergs s}^{-1}$. The spectrum

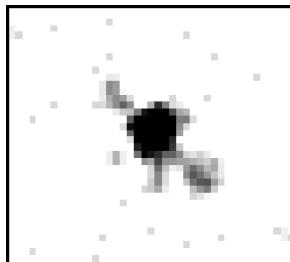


Fig 6.29 A Chandra image of the star DG Tau that shows X-rays from oppositely directed jets: a 25-hour ACIS exposure. The field shown is $20'' \times 17''$. The jets are directed to the upper left and to the lower right. They extend 700 AU from the star (from Güdel et al., 2008).

of the far jet is absorbed more than that of the near jet, perhaps because it is shining through the disc – an intriguing possibility for study of material in the disc.

It seems clear that planets form in discs around young stars. The planets must accrete material and avoid spiraling into the star until the star turns on and blows the gaseous disc away. Turbulence in the disc is reported to increase a planet's chances for survival. Magnetic fields can move ionised material to generate turbulence. Cosmic rays and X-rays can ionise material in the disc. The inner part of the disc shields the disc from X-rays generated close to the stellar surface, but X-rays from a jet, extending far above the disc, can shine directly onto the surface of the disc. So X-rays might play a role in the formation of planets.

6.9 | Summary

The well-observed solar X-rays have led to knowledge about conditions on the solar surface and interior. The same considerations are applied to other stars. It seems clear that coronal plasma is heated and contained by turbulent magnetic fields. The Sun, however, is a weakling. Almost all other stars are X-ray sources, and coronal processes are more extreme. Young stars and some binary systems have $L_x \sim 10^4$ times greater than that of the Sun. This requires some combination of stronger fields, denser plasma and larger confining magnetic structures. Indeed, observations

sometime show coronal extent of several R_* . The number and variety of systems that emit X-rays are extensive, as are the observational techniques used to extract useful information. The next chapter also concerns stars: massive stars, which are also strong X-ray sources but with different production mechanisms.

References

- Ayres, T. R., A. Brown & G. Harper (2003) Buried alive in the coronal graveyard, *ApJ* **598**, 610.
- Brickhouse, N., A. K. Dupree & P. R. Young (2001) X-ray Doppler imaging of 44i Bootis with Chandra, *ApJ* **562**, L75.
- Canizares C., et al. (2000) High-resolution X-ray spectra of Capella: Initial results from the Chandra high-energy transmission grating spectrometer, *ApJ* **535**, L41.
- Güdel, M., et al. (2002) X-ray evidence for flare density variations and continual chromospheric evaporation in Proxima Centauri, *ApJ* **580**, L73.
- Güdel, M., et al. (2008) Discovery of a bipolar X-ray jet from the T Tauri star DG Tauri, *A&A* **478**, 797.
- Haisch, B. M., et al. (1983) Coordinated Einstein and IUE observations of a disparities Brusques type flare event and quiescent emission from Proxima Centauri, *ApJ* **267**, 280.
- Hall, D. S. (1972) A T Tauri-like star in the eclipsing binary RS Canum Venaticorum, *PASP* **84**, 323.
- Ishibashi, K., et al. (2006) Chandra/HETGS observations of the Capella system: The primary as a dominating X-ray source, *ApJ* **644**, L117.
- Kastner, J., et al. (2002) Evidence for accretion: High-resolution X-ray spectroscopy of the classical T Tauri star TW Hydrae, *ApJ* **567**, 434.
- Ness, J.-U., et al. (2002) Coronal density diagnostics with Helium-like triplets: Chandra-LETGS observations of Algol, Capella . . . , *A&A* **394**, 911.
- Ness, J. U., et al. (2003) Modeling the Ne IX triplet spectral region of Capella with the Chandra and XMM-Newton gratings, *ApJ* **598**, 1277.
- Petit, M. (1954) Les étoiles variables du type U. V. Ceti, *Ciel et terre* **70**, 401.
- Preibisch, T., & E. Feigelson (2005) The evolution of X-ray emission in Young stars, *ApJS* **160**, 390.
- Preibisch, T., et al. (2005) The origin of T Tauri X-ray emission: New insights from the Chandra Orion Ultradeep Project, *ApJS* **160**, 401.
- Siarkowski, M., et al. (1996) Corona(e) of AR Lacertae. II. The spatial structure, *ApJ* **473**, 470.
- Stern, R. A., et al. (1992) Ginga observations of X-ray flares on Algol, *ApJ* **400**, 321.
- Vaiana, G. S., et al. (1981) Results from an extensive Einstein stellar survey, *ApJ* **245**, 163.
- Walter, F., P. Charles & S. Bowyer (1978) X-ray emission from UX Ari – RS Canum Venaticorum systems as a class of coronal X-ray sources, *ApJ* **225**, L119.
- BBSO/NJIT (2007) [http://www.bbsr.njt.edu/](http://www.bbsr.njit.edu/).
- JAXA (2007) http://hinode.nao.ac.jp/news_e/20061220_press_e.
- NASA/CXC/M. Weiss (2001) <http://chandra.harvard.edu/photo/2001/0176>.
- NASA/CXC/M. Weiss (2003) <http://chandra.harvard.edu/photo/2003/twhy/>.
- NASA/CXC/Penn State (2005) <http://chandra.harvard.edu/photo/2005/orion/>.
- NASA/SLISR (2009) <http://soi.stanford.edu/results/SolPhys200/Schrijver/TRACEpodarchive.html>.
- NOAA/SWPC (2009a) <http://www.swpc.noaa.gov/today.html>.
- NOAA/SWPC (2009b) <http://www.swpc.noaa.gov/ftpmenu/warehouse.html>.
- SAO (2009) <http://xrt.cfa.harvard.edu/xpow/20070604.html>.
- SIDC Belgium (2009) <http://www.spaceweather.com/glossary/sunspotnumber.html>.
- TRACE (2009) <http://trace.lmsal.com/Science/ScientificResults/>.
- XRT (2009) <http://xrt.cfa.harvard.edu/data/gallery.php>.

Books and reviews

- Getman, K. V., et al. (2005) Chandra Orion ultradeep project: Observations and source lists, *ApJS* **160**, 319.
- Golub, L. & J. M. Pasachoff (1997, Second Edition 2009) *The Solar Corona*, Cambridge: Cambridge University Press.
- Güdel, M. (2006) Physics of stellar coronae, arxiv: astro-ph/0609389.
- Rosner, R., L. Golub & G. S. Vaiana (1985) On stellar X-ray emission, *ARA&A* **23**, 413.

Chapter 7

Early-type stars

7.1 | O stars

The most luminous, most massive stars are the O stars. Starting with more than $25 M_{\odot}$ of material (possibly $\sim 100 M_{\odot}$), they burn their nuclear fuel at a prodigious rate. They live only a short time and end in a brilliant supernova explosion. The surrounding space is left full of stellar debris enriched in heavy elements. Our bodies all contain elements made in these massive stars.

These are not common stars, and none are nearby. The brightest ones visible to the naked eye are δ and ζ Orionis at the two ends of Orion's Belt. Both are 1600 pc distant and spectral type O9.5; ζ Puppis is 2400 pc distant and type O5. Because the nuclear fuel is consumed rapidly, the lifetime is, astronomically speaking, short. In a few million years an O star changes character, becoming perhaps a red giant or a Wolf-Rayet star. We see, with naked eye or telescope, only the outer layer, which gives little information about events in the core. Hidden from view, the central region evolves rapidly until the nuclear fuel is exhausted. As explained in Chapter 8 on supernova remnants, the core collapses and the gravitational energy released powers the resulting supernova. That is the end of the O star.

Astronomers originally believed all stars evolved along the main sequence. In this scheme a star would start life as a hot O star and, as it aged, would change into progressively cooler spectral types. The O stars and B stars were thus called 'early' stars. Although this idea has long been abandoned, the nomenclature lives

on! Hot stars are 'early type' and cool stars are 'late type'.

O stars are not scattered randomly throughout the galaxy. Most are found in rather loose groups or clusters called 'OB associations' which delineate the spiral arms of our Galaxy. They are usually close to clouds of dust, nebular variables and other indications that star formation is in progress. This is expected. Since O stars do not have long lives, they do not usually drift far from their birthplace.

The energy output of O stars is enormous. A star of spectral type O5I has a luminosity 5×10^5 that of the Sun. If such a star were as close to us as Arcturus, a distance of 10 pc, it would have a visual magnitude of -7 and would appear 20 times brighter than Venus at its brightest! Because the surface temperature is 40 000 K, most of the energy is radiated in the UV range. Only 4 per cent of the energy is found in the visible band. A lot of energy is also expended in the ejection of a hot stellar wind. A mass loss of $10^{-6} M_{\odot} \text{ yr}^{-1}$ at a terminal velocity of 2000 km s⁻¹ is not unusual. The wind from the Sun is a gentle breeze compared with the hurricane in space surrounding an early-type star.

O stars have a profound effect on their surroundings. After the star is formed, the energy generated by the star soon ionises and sweeps away any material remaining in the vicinity. The strong stellar wind pushes matter outward. There should be a central region around the star with wind flowing freely, a surrounding volume of hot (perhaps X-ray emitting) gas and a relatively cool and dense outer shell. This shell contains most of

the mass and should expand slowly into the ISM. A ‘bubble’ is thus formed around the star and is predicted to have a diameter of 10–30 pc. A few such bubbles have been discovered optically, but the geometry is usually more complex than a simple sphere.

7.2 | X-rays

All O stars are X-ray sources and consistently radiate $\approx 10^{-7}$ of their bolometric luminosity as soft X-rays. These X-rays are produced in the strong outflowing wind, but the mechanism is not well understood. One of the most luminous O stars in our Galaxy is HD93129A, located in the Carina Nebula. The bolometric luminosity of this star is 2×10^{40} ergs s $^{-1}$. The energy radiated as X-rays, 2×10^{33} ergs s $^{-1}$, although an insignificant part of the output of this star, is comparable to the total output of the Sun at all wavelengths!

The first detection of X-rays from O stars occurred on 15 December 1978. On this occasion, 3 weeks after launch, the Einstein telescope was pointed at the bright X-ray binary Cyg X-3. This source was to be used to calibrate the IPC detector. The appearance of O stars was an unexpected result, a thrill for the scientists examining these first pictures but a shock for Rick Harnden, the person responsible for the calibration of this detector.

This was the first operation of the detector in space, at the focus of the telescope and pointed at a bright X-ray source. There had been 2 years of laboratory calibrations. All the variable instrument parameters had been set to give optimum and reliable operation. Here was the first opportunity to observe the detector response to a bright cosmic source. The source, Cyg X-3, was clearly visible in the field and the spatial spread caused by instrumental effects was about as expected, but it was with sinking heart that Harnden first contemplated the confusion in the centre of the field (Fig. 7.1).

Several irregular bright spots (looking particularly abnormal when viewed with the quick-look TV monitor at the control centre) appeared in a region of the sky thought to be free of sources. An anxious 3 days passed, filled with thoughts

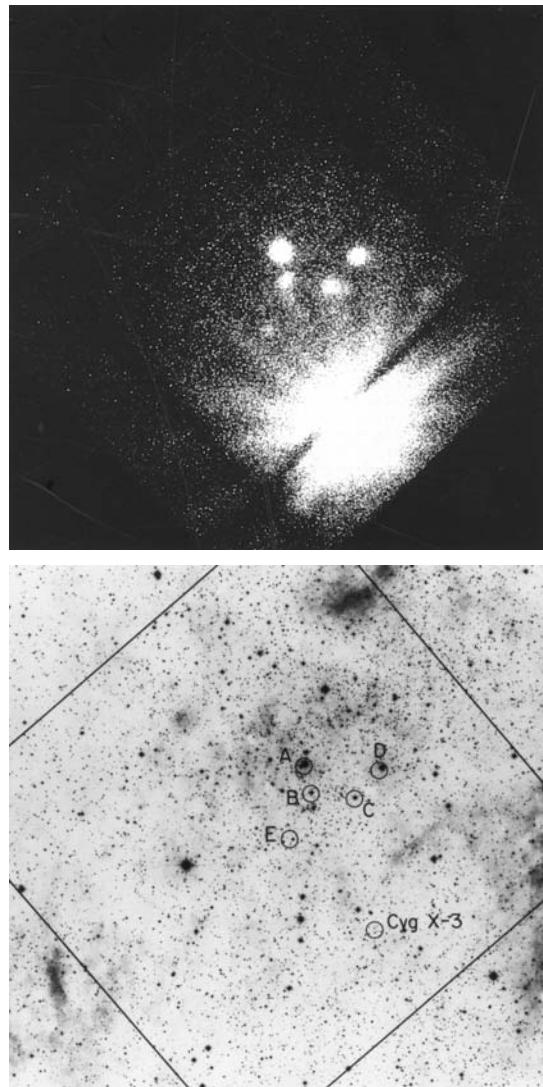


Fig 7.1 The first O association discovered to emit X-rays. (top) A 3-hour Einstein IPC observation. The field is 1° square shadowed by four window-support ribs, making a tic-tac-toe pattern. The image is printed to show the faint sources at the centre. The lower field is filled with a ragged halo surrounding the very-bright, overexposed Cyg X-3 image. This halo is due to imperfections in the telescope mirror, and the rays are from the mirror-support structure. The weaker X-ray sources are bright O stars in the Cyg OB2 association. (bottom) The Cyg OB2 association as seen in the red Palomar Sky Survey. Circles show stars which are prominent X-ray sources, and the box indicates the boundary of the IPC field of view (from Harnden et al., 1979).

of ‘hot spots’ caused by dirt on the wires within the proportional counter (and impossible to clean at this point). All ended happily with the realisation that the detector was operating properly and that this emission was real. When the X-ray picture was overlaid on the Palomar Sky Survey, there was a bright star at the position of each X-ray bright spot in the centre of the field. The pattern of stars was the same as the pattern of X-ray spots, an unmistakable identification. Thus O stars were found to be strong X-ray sources. These particular stars are all members of the highly-reddened Cyg OB2 association. It is 2 kpc distant, and X-rays are attenuated by a factor of 5–10 in their journey to the vicinity of Earth. Only because the stars are so very luminous were the X-rays detected.

7.3 | Stellar winds

One of the most interesting developments in UV astronomy was the observation of strong stellar winds. Theories of stellar evolution which assumed the mass of the star remained constant had to be modified. Some massive early stars lose 10^{-6} – $10^{-5} M_{\odot}$ yr $^{-1}$ (M_{\odot} = 1 solar mass). The outflowing wind absorbs radiation emitted from the stellar surface, and this accelerates material to escape velocity. The material is ‘warm’ and contains ions which have strong characteristic absorption lines in the UV. The strength of each absorption line is a measure of the amount of material in the wind, and the Doppler-shifted structure of the line is a measure of the velocity of the material. Lines of highly ionised carbon, nitrogen and silicon serve as good wind indicators. Winds from the hottest O stars have terminal velocities of 3500 km s $^{-1}$, three times the escape velocity. Stars with highest mass-loss rates are the Wolf-Rayet stars with winds carrying over $10^{-5} M_{\odot}$ yr $^{-1}$ away from the star. These may be compared with the relatively feeble solar wind with velocity 5–10 times less and mass loss a factor of 10^7 less.

Figure 7.2 shows an early-type star UV spectrum dominated by two broad absorption features and containing many narrow, weak absorption

Box 7.1 | The solar wind

The hot solar atmosphere, the corona, contains thermal particles with energies up to a few keV. Particles escape from the outermost corona and stream away from the Sun. This outflow is called the *solar wind* and contains protons, He nuclei, electrons and some heavy ions. In the vicinity of Earth typical parameters are: density, 5 particles cm $^{-3}$; velocity, 400 km s $^{-1}$. There is considerable variation because of solar activity, but on average the energy in the wind is $\sim 10^{-7} L_{\odot}$ and the mass loss is $\sim 10^{-14} M_{\odot}$ yr $^{-1}$. The kinetic energy of a proton is ~ 1 keV and that of an oxygen ion is ~ 16 keV. The solar wind pushes the tails of comets away from the Sun and produces X-rays from collisions with cometary material. As a matter of interest, magnetic field is carried away from the Sun embedded in the solar wind. In the vicinity of the Earth, this has a value $\sim 5 \times 10^{-5}$ G. The geometry of Earth’s magnetic field is shown in Fig. 1.2.

lines. Absorption of the stellar continuum by stationary material at the base of the wind produces the narrow absorption features. The strong broad double absorption line is because of the ion Si IV in the outward moving wind. The normal position of this feature is at 1393 and 1402 Å. Line centroids are Doppler shifted ≈ 3 Å to shorter wavelengths

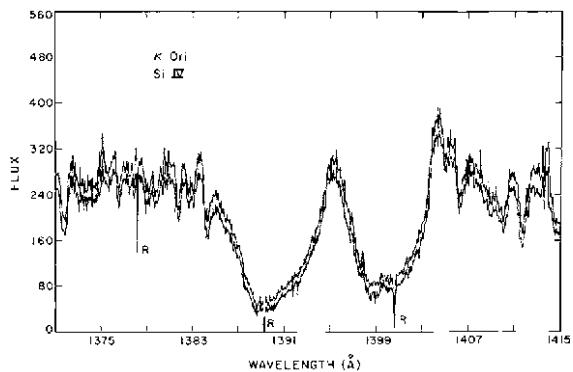


Fig 7.2 UV spectrum of the second-magnitude star κ Ori, a B supergiant, taken with the IUE satellite. The shape of the two broad absorption lines is determined by characteristics of the stellar wind. The features labelled ‘R’ are calibration marks (courtesy of A. Dupree).

Table 7.1 | X-rays and winds from some bright early stars.

Star	Distance (kpc)	Magnitude (V)	Spectral type	v_∞ (km s $^{-1}$)	\dot{M} (M $_{\odot}$ yr $^{-1}$)	L_X (erg s $^{-1}$)
ζ Pup	0.44	2.3	O4If	2200	3×10^{-6}	2.6×10^{32}
δ Ori	0.5	2.2	O9.5II	2000	1×10^{-6}	3.7×10^{32}
θ^1 Ori C	0.5	5.1	O6V	2500	4×10^{-7}	5.1×10^{32}
ζ Ori	0.5	1.6	O9.7Ib	2220	1×10^{-6}	5.0×10^{32}
Cyg OB2 8A	1.8	9.0	O5.5I(f)	2650	1.3×10^{-5}	1.3×10^{34}
WR 140 A	1.4	6.8	WC 7	2860	6×10^{-5}	2×10^{34}
η Car ¹	2.6	–1 to 7	LBV	500/3000 ¹	$3 \times 10^{-4}/10^{-5}$	$3\text{--}6 \times 10^{33}$
HD 93129 AB	2.8	7.3, 8.9	O2I, O3V	3200	1.8×10^{-5}	4×10^{33}

¹ a binary system.

because the material is moving towards us. The rather sharp inflection shows the maximum velocity of the bulk of the wind. Wind characteristics of some of the stars discussed in this chapter are given in Table 7.1.

7.4 | X-rays from single stars

The O-star wind is accelerated by absorption of UV radiation from the stellar surface. The temperature is high and outward velocity is low close to the surface. Ions cool as the wind expands outward and velocity increases to terminal velocity v_∞ . Spectral lines exhibit both thermal and Doppler broadening. The latter is dependent on magnitude of velocity and the range of angles between velocity vectors and the line of sight. If X-ray emission occurs close to the surface, Doppler broadening will be small, and the stellar disc will occult material moving directly away from the observer. If emission arises far from the surface, velocity is higher and Doppler broadening will be larger. If the wind is transparent, lines will be symmetrical about the rest wavelength.

A UV pressure-driven wind is not stable, and shocks are expected to develop. The basic model assumes many fragmentary shocks distributed throughout the wind. The wind is massive, absorption is strong and observed X-rays cannot come from the vicinity of the surface of the star. Observations with low-resolution spectrometers

on Einstein, ROSAT and ASCA allowed calculation of absorption, temperature, emission measure and variability. The average temperature was found to be ≈ 0.5 keV, but spectra were not well fitted by single-temperature models. The emission measures showed that only a small fraction of the wind material was generating X-rays. Variability in O-star X-ray emission was rare. High-resolution spectrometers on Chandra and XMM offered a chance to determine where, within the wind, the X-ray lines were generated. As usual, results vary from star to star. Because the distributed-shock model does not explain the lack of variation of emission from most single stars nor the observed almost-constant ratio L_X/L_{bol} , this model is not expected to apply to all cases.

Figure 7.3 is a compilation of Chandra spectra showing X-ray emission lines of several O stars. Higher temperatures result in more ionisation and more energetic lines. Note that the line shapes vary. δ Ori, for example, has a narrow 12 Å FeXVII line compared to the other stars.

Figure 7.4 shows line shapes measured from a 19-hour observation of ζ Pup and a 14-hour observation of δ Ori. Emission lines of ζ Pup have characteristics close to expectations of the distributed-shock model (Cassinelli *et al.*, 2001). Four lines are shown in Fig. 7.4. The NvII line is flat topped and symmetrical about the rest wavelength, as would be the case for emission from a fast-moving transparent shell. Other line shapes are skewed towards shorter wavelengths. As expected, absorption in the wind is

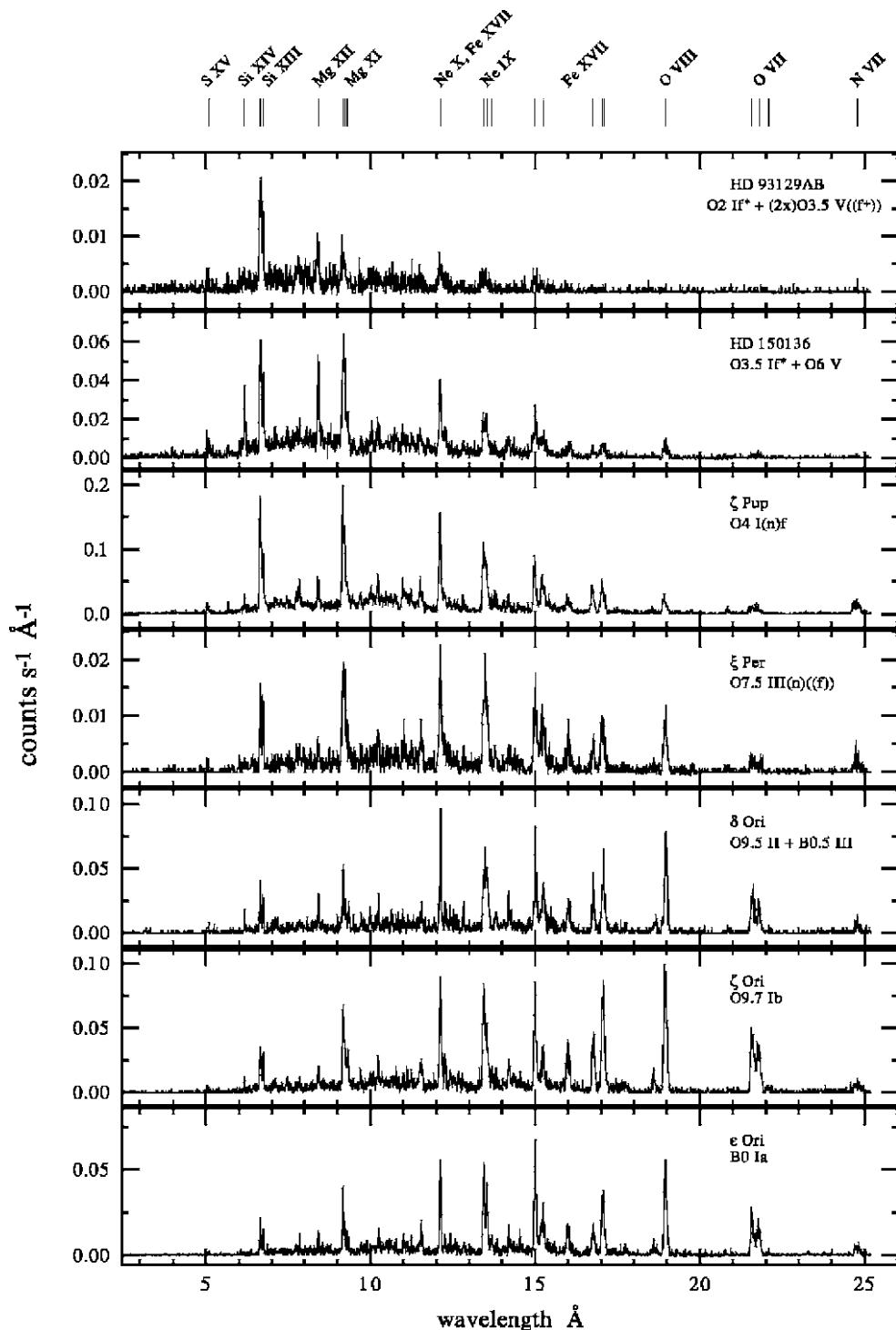


Fig 7.3 Emission lines from seven main sequence and giant O stars observed with the Chandra high-resolution spectrometer. Stellar surface temperature decreases going from top to bottom. Emitting ions are identified along the top boundary (courtesy of W. Waldron; Waldron & Cassinelli, 2007).

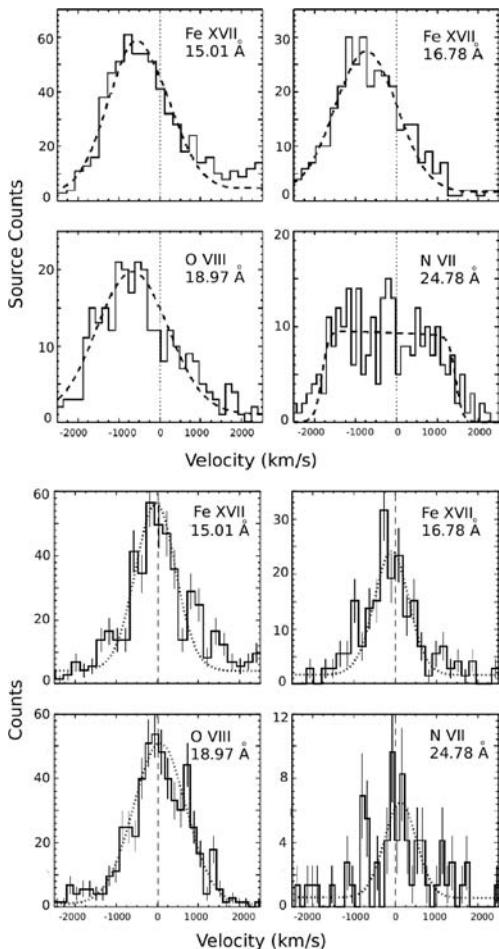


Fig 7.4 Emission line shapes measured by the Chandra high-resolution grating. Vertical markers indicate rest wavelengths of the transitions. At top, four lines from ζ Pup are broad and Doppler shifted, as expected for lines generated in a wind with some absorption. At bottom, the same four lines from δ Ori show less Doppler broadening and no bulk motion (from Cassinelli *et al.*, 2001; Miller *et al.*, 2001).

preventing red-shifted emission from the back side from reaching the observer. Doppler broadening is $\approx 600\text{--}1500 \text{ km s}^{-1}$, close to the terminal velocity of 2000 km s^{-1} .

On the other hand, lines from δ Ori are narrow and unshifted. Doppler broadening is only $\approx 400 \text{ km s}^{-1}$, 20 per cent of v_∞ , contrary to prediction of the distributed-shock model (Miller *et al.*, 2001). The assumption of a great many small shocks might account for the lack of X-ray

variability and also produce a strongly clumped structure, transparent to far-side emission and with blue shift absent.

Lines from the wind of θ^1 Ori C also do not fit the model. This unusual O star has a strong magnetic field which probably affects X-ray production. The emitting plasma is apparently produced in the equatorial region, and temperatures are higher there than expected from the wind-shock model. Lines are narrow, and there is some modulation of flux from rotation of the star (Gagne *et al.*, 2005). So deviations from the distributed-shock model have been attributed to a clumpy wind and to magnetic structure. To bring the discussion full circle, we note that the Cyg OB2 association in Fig. 7.1 has been observed with Chandra, and the lines also show smaller shifts and asymmetries than expected. Much work remains to be done before O-star emission is understood.

7.5 Colliding winds

Many O stars are found in binary systems, which contain two early stars, each with a strong wind. The winds, initially directed outwards from each star, collide, and copious X-rays can come from the resulting shocked gas. Two of the brightest examples are WR 140 in the constellation Cygnus and the peculiar object η Carinae. Colliding winds are expected to produce a harder spectrum and a higher L_X/L_{bol} ratio than emission from single stars.

WR 140 is a sixth-magnitude binary consisting of an O star (O4) and a Wolf-Rayet star (WC 7) in a 2899-day (8-year) highly eccentric orbit. X-ray emission varies and is maximum around periastron, when the two stars are closest. The X-ray flux, monitored with RXTE, reaches maximum a month before periastron and then drops a factor of 10 because of absorption in the cool, dense wind of the WC 7 star. Because the WC 7 wind is the more energetic, the interface between the two winds is pushed into a roughly conical shape around the O4 star, as shown in Fig. 7.5. The X-rays should come largely from the region directly between the two stars where the density of shocked gas is highest. Emission is highest at periastron because

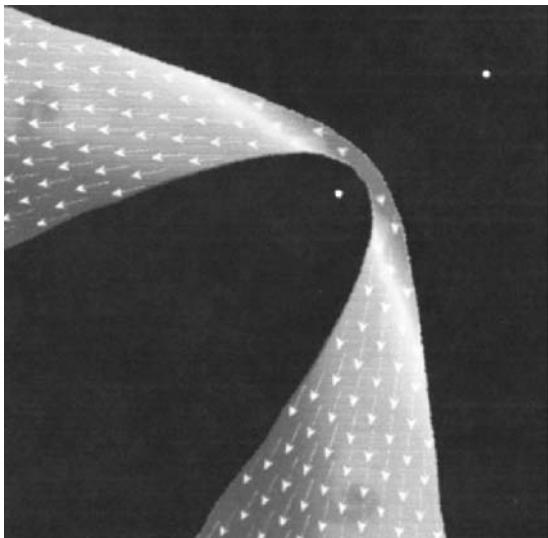


Fig 7.5 Colliding wind geometry of the binary star WR 140. Dark regions do not emit radiation. They are filled with the cold outflowing winds from the two stars, both shown as white dots. The gas is shock heated where the winds collide and then streams outward, as shown by arrows. At X-ray maximum the orientation of the binary corresponds to viewing from the bottom of this diagram, so the line of sight to the area of maximum emission is largely through the ionised outflowing wind (from Pollock et al., 2005).

density of the shocked wind in this region is highest then. Soon after periastron, this structure has rotated so the emitting region is viewed through the dense, cold (because it has not passed through

the shock) wind of the WC 7 star, and the softer X-rays are absorbed.

Figure 7.6 shows the Chandra high-resolution spectrum of WR 140 taken at maximum X-ray brightness. About 70 emission lines are superposed on a continuum, which accounts for 80 per cent of the emission. The continuum spectrum is thermal with temperature $kT \approx 5$ keV and with absorption in the ISM and in circumstellar material. Centroids of the lines are all blue shifted because the ions are in the shocked wind streaming towards the observer at $400\text{--}700$ km s^{-1} . The equilibrium of the shocked material is complicated. There is some evidence that the shock is ‘collisionless’. Cold wind material may be ionised through interaction with high-velocity protons, not by collision with the hot electrons which produce the continuum. Shock conditions are similar to those produced when supernova ejecta plow into circumstellar material. Chapter 8 has more information on collisionless shocks and non-equilibrium plasma.

7.6 | Eta Carinae

X-rays from colliding winds are also produced by η Carinae, which, for years, has been a unique object of great interest. The star η Car lies in the middle of the brightest part of the Carina Nebula and was

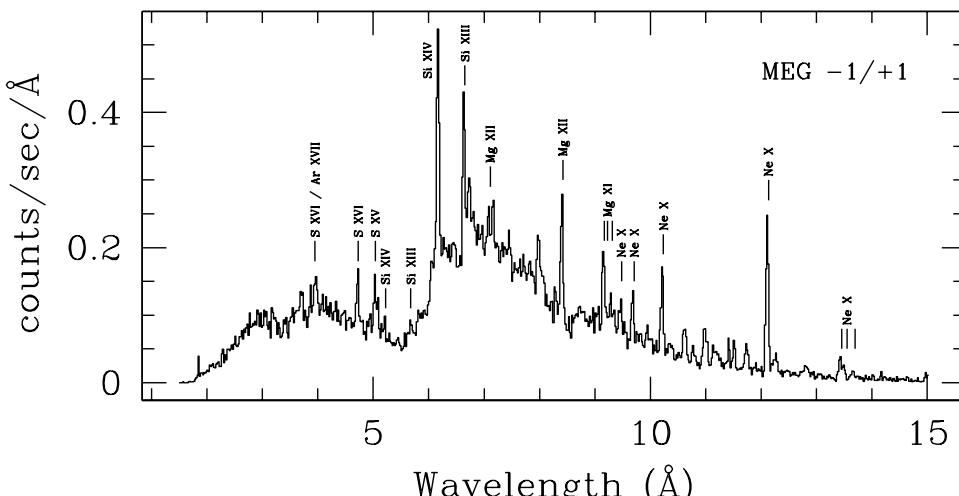


Fig 7.6 A high-resolution Chandra grating spectrum of WR 140 taken close to maximum X-ray emission. Note that in comparison with the spectra of Fig. 7.3, the temperature is high and the continuum strong (from Henley et al., 2004).

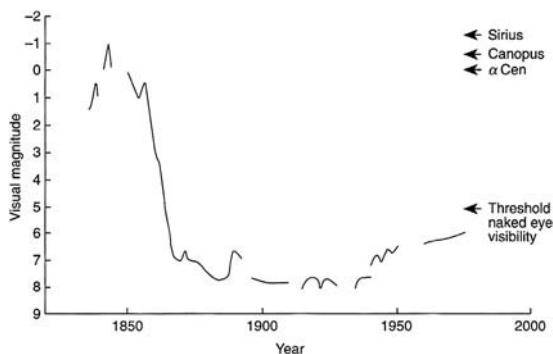


Fig 7.7 The light curve of η Carinae since 1836. It was brighter than first magnitude for more than 20 years and has remained just below the threshold of naked-eye visibility for more than a century.

first noticed as a bright, variable star. A light curve spanning 150 years is shown in Fig. 7.7. Initially of fourth to second magnitude, the star brightened to first magnitude in the 1830s and reached a maximum brightness of -1 in 1843. It remained above first magnitude until 1857, declined steadily from first to seventh magnitude between 1857 and 1869 and has varied irregularly over the last 100 years between sixth and eighth magnitude.

Star η Carinae forms the bright core of a small optical nebula, the *homunculus* (meaning ‘little man’ because such was its apparent shape). This $12'' \times 17''$ nebula was resolved by the Hubble Space Telescope into a bipolar structure (Colour Plate 16). The structure is expanding with a velocity of about 500 km s^{-1} , consistent with the hypothesis that this material was ejected during an 1843 or later outburst. The homunculus is in turn surrounded by a faint $22'' \times 45''$ outer shell comprising discrete optical features and expanding with velocities of $300\text{--}1300 \text{ km s}^{-1}$. A few other LBV stars, the Luminous Blue Variables, also show evidence for past great eruptions in which a few M_{\odot} of material were ejected with kinetic energies of $\sim 10^{49}\text{--}10^{50} \text{ ergs}$.

In the infrared, η Carinae is spectacular. At a wavelength of $20 \mu\text{m}$, the homunculus is the brightest extrasolar infrared source in the sky. At a distance of 2600 pc , the bolometric luminosity is $6 \times 10^6 L_{\odot}$. The source of this emission is a shell of dust surrounding a central massive star. The dust absorbs optical and UV energy radiated by

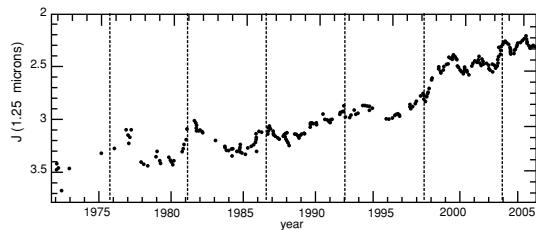


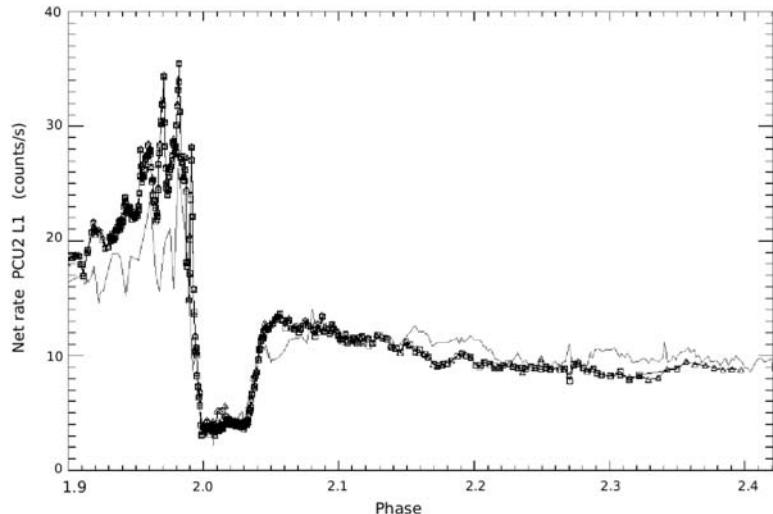
Fig 7.8 A 35-year record of infrared emission from η Carinae from SAAO. The J-band magnitude is plotted as a function of time. Vertical lines indicate the 5.5-year period, but no clear periodic effects are seen (courtesy PAC).

the star and re-radiates it as infrared. Furthermore, optical emission from the homunculus is strongly polarised, as expected of light scattered from dust. The luminous central object is now known to be a massive binary, but the dust shell is thick enough to hinder direct optical observation. It has been suggested that this massive star may well become the next galactic supernova. The great luminosity certainly implies a very short lifetime. The infrared emission has been monitored since 1972 by SAAO and, as shown in Fig. 7.8, has been steadily increasing.

X-rays with energies above $\approx 2 \text{ keV}$ do penetrate the dust shell, and the emission of the central star has been monitored by RXTE for over 11 years. There is optical evidence for a period of 5.5 years, and the X-ray light curve so far has shown three sharp, eclipselike minima repeating at this interval (Fig. 7.9). The X-ray emission is thought to be from colliding winds, and the shape of the light curve implies an eccentric orbit similar to that of WR 140 (Ishibashi *et al.* 1999).

Star η Carinae itself is an X-ray source but not exceptionally strong. The X-ray luminosity is less than 10^{-6} of the total luminosity. The Chandra ACIS observation is shown in Colour Plate 16. The binary star appears as a central, pointlike, strongly absorbed source. This is surrounded by a ring of soft emission, which comes from the optical ‘outer shell’. The X-ray bright spots are not far from bright optical knots. In analogy with supernova remnants, the soft X-ray emission comes from an expanding shell interacting with surrounding material, as deduced from optical data. Star η Car is but one member of a spectacular grouping of early stars in the Carina Nebula.

Fig 7.9 The hard X-ray light curve of η Carinae monitored by RXTE as a function of phase. One half cycle, starting just before the beginning of the second observed minimum, is shown. Data points give intensity through the July 2003 minimum. The faint line traces data through the previous minimum in December 1997 (from Corcoran, 2006).



7.7 O stars in the Carina Nebula

The Carina Nebula encompasses several O associations and shows why O associations are interesting targets for an X-ray telescope. Closely packed pointlike sources are easily separated from one another and from surrounding diffuse emission.

Located in the far southern sky, it is one of the brightest optical emission nebulae in the Milky Way. The nebula subtends about 1 square degree and can easily be seen with the naked eye. An optical photo (Colour Plate 17) shows detailed structure within the nebula and many bright O stars (of magnitude 6–9). This nebula contains a collection of the hottest and most massive stars known, many of spectral type O3. There are also three Wolf-Rayet stars and η Car itself, perhaps the most luminous star in our Galaxy. Two prominent dust lanes give an easily recognisable dark 'V' just below and on either side of the brightest region. The nebula contains shells of rapidly moving gas and dust which are illuminated by UV radiation from early stars and re-radiate much of this energy as optical emission lines. Here is an area in which massive stars have just been, and are still being, created.

This field is a nice illustration of the power of the X-ray telescope. X-rays from this entire

complex were first detected in the 1970 rocket scan shown in Fig. 1.6. A small maximum labelled ' η ' in the middle energy band indicates a barely detectable source. The identification as η Car was a guess, a mention of an unusual source in a very large error box. The 1980 Einstein X-ray picture of the Carina Nebula is shown in Fig. 7.10. Not only was η Car itself detected but also strong diffuse X-ray emission extending over most of the nebula. This diffuse emission accounts for 80 per cent of the X-rays and the 1970 detection. All of the most luminous O stars show as unresolved X-ray sources. The 2005 Chandra image of one of these Einstein sources, the cluster Tr 14, is seen in Colour Plate 18, and the cluster is resolved into dozens of pointlike sources. Thus one of the weakest sources known in 1970 is now resolved into hundreds of stellar sources and a large region of diffuse emission. The star η Car itself is indeed a source and has complex morphology and variability.

The optical-emitting diffuse gas in the Carina Nebula is mostly hydrogen and fills a region about 50 pc in diameter. This is the stuff from which stars are being made. There are $\sim 3 \times 10^5 M_\odot$ of visible matter. The dark clouds are presumably much denser and might contain at least this much material. An estimate of $10^6 M_\odot$ for the mass of the nebula is not unreasonable.

The diffuse hot gas which emits X-rays is interspersed with this cooler material. The hot gas and



Fig 7.10 | Contours from Einstein observations of the Carina Nebula overlaid on the optical image. Four 2-hour IPC observations have been combined. The easternmost source (at the far left) is an anomalous X-ray Pulsar (see Chapter 9) with period 6.4 s and is not associated with the Carina Nebula. The three bright sources at the centre are η Car (lower left); a Wolf-Rayet star, HD93162 (lower right); and an O star, HD93129 (upper right). All other sources are also O stars. There is also diffuse emission, probably powered by stellar winds.

the cool gas are not mixed uniformly but co-exist in irregular filaments and patches. The pressure of a cloud of hot gas is roughly equal to that within the denser, cooler material. The brightest regions of diffuse X-ray emission have a size of about 1 pc and a density of about 1 atom cm^{-3} . The size and density of the hot regions, however, vary greatly throughout the nebula, and only about $\sim 30 M_{\odot}$ of hot gas are required to account for the diffuse X-ray emission. This multiphase structure is found throughout the ISM and is discussed in Chapter 5.

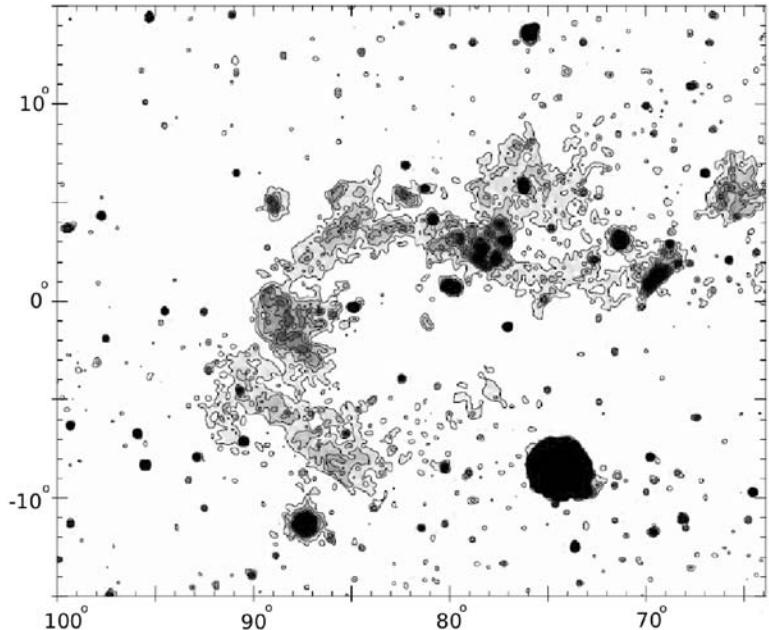
The energy and mass needed for this diffuse hot gas are easily supplied by the winds of the early-type stars. They soon create high-pressure regions which change the morphology of the cool clouds from which they formed. Thus the expected ‘bubbles’ around O stars exist, but because the star-forming cloud is dense and irregular, the appearance is not simple. Only if the star were created in a uniform medium would a large, spherical bubble be formed by the wind. Figures 7.10 and Colour Plate 18, then, are observations of hot gas formed far from the stars by the energy carried in stellar winds. This will eventually displace the cooler star-forming material from a large volume surrounding the cluster.

7.8 | Superbubbles

A superbubble is a large cavity in the ISM caused by deposition of energy from supernovae and O stars. Generally, there is a shell of H (emitting optical H α and radio), an interior filled with hot gas ($T \approx 10^6 \text{ K}$ emitting X-rays) and a diameter of 100 pc or more. Superbubbles are readily detectable in external galaxies where soft X-rays are not absorbed by gas in the line of sight. Many have been found in the Magellanic Clouds, where they were spatially resolved by ROSAT (e.g. Dunne *et al.*, 2001). In our Galaxy, absorption by intervening gas and dust in the ISM makes identification difficult.

A possible galactic superbubble, more than 400 pc in diameter, was discovered in Cygnus in 1980 through its X-ray emission. If real, it would be one of the largest known objects in our Galaxy. (In astronomy the term *object* embraces many phenomena, usually referring to a spatially isolated source of radiation at a particular wavelength.) It was found using a set of large-area soft X-ray detectors designed to map the soft X-ray background. During the analysis of the HEAO-1 scans of the sky, Cash and Charles (1980) found a previously uncatalogued source in Cygnus which was of low

Fig 7.11 A ROSAT map of the Cygnus superbubble at 1.5 keV. Coordinates are galactic. The overexposed Cygnus Loop is at (74, -08) and is closer than the superbubble. The bright pointlike sources Cyg X-1, Cyg X-2 and Cyg X-3 are at (71, 03), (87, -11) and (80, 01), respectively. X-rays from the interior are absorbed by material in the Cygnus Rift which lies in the galactic plane running horizontally across the figure (from Uyaniker *et al.*, 2001).



temperature and obviously extended (Cash *et al.*, 1980). It looked like a supernova remnant, but when adjacent 3°-wide scans were examined, it was seen to be part of a larger structure. Ten years later, this region was also mapped by ROSAT, as shown in Fig. 7.11.

This map encompasses most of Cygnus and shows several other well-known X-ray sources, including Cyg X-1, Cyg X-2, the Cygnus Loop and a lesser-known supernova remnant, G 65.2+5.7. The feature of interest is a partial ring of X-ray emission 13° by 18° in extent. The spectrum of this huge object indicates a temperature of 2 million degrees. The observed morphology is shaped by absorption in cold interstellar material. In particular, gas associated with a large dust lane in Cygnus (the Cygnus Rift) absorbs all the X-rays from the centre of the structure. This absorption in the galactic plane gives the superbubble a horseshoe-shaped appearance.

The Cygnus dust lane is about 1.5 kpc distant, so the distance of the X-ray bubble is at least 2 kpc. At that distance, the ring is 400 pc in diameter – 10 times the size of the Carina Nebula and 14 times the size of the Cygnus Loop. Even though the density is low, more than 10^{52} ergs of thermal energy (20 times that of an old supernova remnant) are locked up in the hot gas. The key to understanding the superbubble is the source of this energy.

There are an exceptional number of young, massive stars in the region, including the heavily reddened Cyg OB2 association near the centre of the ring and shown in Fig. 7.1. Despite the high luminosities of these O stars, there are not enough to supply the energy requirements of the superbubble.

Although an O5 star might radiate 10^{53} ergs of UV energy during its lifetime, most of this is absorbed by hydrogen and re-radiated in the optical band. In any case, this UV energy cannot heat dilute gas above 10^5 K. The stellar wind, however, can deposit 10^{50} ergs in surrounding material over the lifetime of the star, and this energy can heat the interstellar material to a temperature of 10^6 K. Finally, some, perhaps all, early stars will eventually explode as type II supernovae, and each will deposit an additional 10^{50} ergs in the ISM. It appears then that 100 O stars could supply the energy for the superbubble through their winds and through supernova explosions.

It is unlikely that such a large structure was created by a single explosion. A super supernova, releasing 10^{54} ergs, could create a large shell, but X-ray emission 100 times brighter than observed would be expected. The bubble was probably built by a chain of ‘normal’ supernova explosions all occurring in a localised region of space but exploding at intervals throughout the life of the bubble.

The observed size and velocity of expansion show the Cygnus superbubble to be at least 3 million years old. It could not have been created by stars in the presently observed Cyg OB2 association, which is only about 1 million years old. It is more reasonable that both the superbubble and the Cyg OB2 association were created by an earlier generation of O stars which have since exploded and vanished. A possible sequence of events leading to the superbubble as we see it today is shown in Fig. 7.12.

Because supernovae produce shock waves which compress material in the ISM, the superbubble is a catalyst to star formation as well as a side effect. Known star-forming sites may well contain evidence for other superbubbles. However, the sheer size of superbubbles and the usually high absorption in the plane of our Galaxy make it difficult to identify them by their X-ray emission alone. Indeed, Uyaniker *et al.* (2001) argue that the Cygnus superbubble is an illusion – a chance superposition of emission regions at different distances along a spiral arm. Perhaps the most promising other candidate is the Orion-Eridanus X-ray hot spot, which extends 30° across the sky and is associated with a large arc of nebulosity known as Barnard's Loop, which envelops most of Orion. At the eastern edge of this region is a large system of gas and dust which includes the very well known Orion and Horsehead nebulae (see Chapter 6). Nearby are the bright O and B stars of Orion's Belt and Sword. Although 21-cm radio emission is notably lacking, there is little doubt that we are seeing a smaller version of a superbubble (Guo *et al.*, 1995).

The process by which a superbubble causes (and is caused by) a chain reaction of star formation could be very effective and may produce most of the new stars in the Galaxy. Indeed, the Solar System might have been created at the edge of a superbubble 4.5 billion years ago.

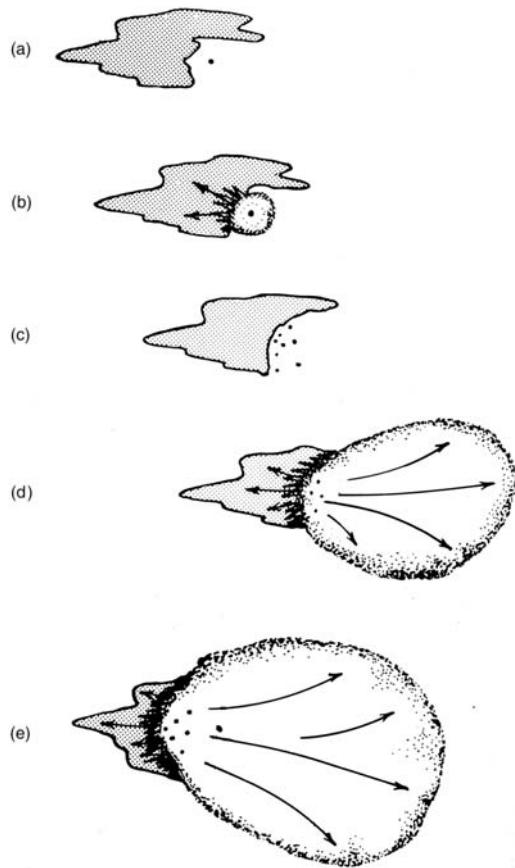


Fig 7.12 The formation of a superbubble. (a) A massive star evolves rapidly and explodes (b) as a supernova, driving a shock wave into a nearby large interstellar cloud and (c) precipitating the formation of new stars. (d) As these stars in turn evolve and explode, their combined blast waves blow a large, hot bubble in space and also eat farther into the cloud, causing more star formation. (e) The process continues as the sputtering supernova wind from the next generation of stars inflates the bubble into a superbubble.

schedule and was pleased that the show would cover a topic of particular interest to him. He thought the show was starting rather obliquely with a long section about sailboats in the middle of the Atlantic Ocean. He soon learned that to a yachtsman, and to the producers of that program, the letters 'OSTAR' meant 'Observer Single-handed TransAtlantic Race', a heroic test of sailor and ship sponsored by the British newspaper *The Observer*. The programme was fascinating, but when an astronomer sees these five letters, a quite different phenomenon springs to mind. We hope, dear

7.9 The astronomer views the rest of the world

A few years ago, F.D.S. sat down to watch what he thought was a television programme about astronomy. He was surprised to find O stars listed in the

reader, that you will now share this enthusiasm for early-type stars.

References

Observations

- Cash, W., et al. (1980) The X-ray superbubble in Cygnus, *ApJ* **238**, L71.
- Cassinelli, J. P., et al. (2001) Chandra detection of Doppler-shifted X-ray line profiles from the wind of ζ Puppis (O4 f), *ApJ* **554**, L55.
- Corcoran, M., (2006) http://asd.gsfc.nasa.gov/Michael_Corcoran/eta-car/etacar-rxte-lightcurve/index.html.
- Dunne, B. C., S. D. Points & Y-H. Chu (2001) X-Rays from superbubbles in the large Magellanic Cloud. VI. A sample of thirteen superbubbles, *ApJ Supp* **136**, 119.
- Gagne, M., et al. (2005) Chandra HETGS multiphase spectroscopy of the young magnetic O star θ^1 Orionis C, *ApJ* **628**, 1005.
- Guo, Z., et al. (1995) X-ray morphology, kinematics, and geometry of the Eridanus soft X-ray enhancement, *ApJ* **453**, 256.
- Harnden, F. R., Jr, et al. (1979) Discovery of an X-ray star association in VI Cygni (Cyg OB2), *ApJ* **234**, L51.
- Henley, D. B., et al. (2004) Probing colliding wind binaries with high-resolution X-ray spectra, *X-ray Radio Connect*, <http://www.aoc.nrao.edu/events/xradio>.
- Ishibashi, K., et al. (1995) Recurrent X-ray emission variations of η Carinae and the binary hypothesis, *ApJ* **524**, 983.
- Miller, N. A., et al. (2001) New challenges for wind shock models: The Chandra spectrum of the hot star δ Orionis, *ApJ* **577**, 951.
- NASA/CXC/GSFC (2007) <http://chandra.harvard.edu/photo/2007/etacar/>.
- NOAO/AURA/NSF (1975) <http://www.noao.edu/image-gallery/html/im0061.html>.
- Pollock, A. M. T., et al. (2005) Bulk velocities, chemical composition, and ionization structure of the X-ray shocks in WR 140 near Periastron as revealed by the chandra gratings, *ApJ* **629**, 482.
- Townsley, L. K., et al. (2005) Parsec-scale X-ray flows in high-mass star-forming regions, *IAUS* **227**, 297.
- Uyaniker, B., et al. (2001) The Cygnus superbubble revisited, *A&A* **371**, 675.
- Waldron, W. & J. P. Cassinelli (2007) An extensive collection of stellar wind X-ray source region emission line parameters, temperatures, velocities, and their radial distributions as obtained from Chandra observations of 17 OB stars, *ApJ* **668**, 456.

Catalogues, reviews, popular

- Cash, W. & P. Charles (1980) Stalking the Cygnus superbubble, *Sky Telescope* **59**, 455.
- Chlebowski, T., F. R. Harnden Jr & S. Sciortino (1989) The Einstein X-ray Observatory catalog of O-type stars, *ApJ* **341**, 427.
- Rauw, G., (2006) X-ray emission from early stars: New results and new challenges, in *The X-ray Universe 2005*, El Escorial, Madrid, Spain.

Chapter 8

Supernova explosions and their remnants

8.1 | Introduction

This chapter describes phenomena caused by truly large explosions: catastrophic events in which large stars disintegrate completely. Vast clouds of stellar debris are ejected and are rapidly heated to temperatures of millions of degrees. This is a very important mechanism in astronomy, as it enriches the ISM with heavy elements, out of which new stars and planetary systems (such as our own) can be formed. These expanding clouds of hot gas are strong sources of X-ray and radio radiation. They shine clearly as extended objects with a great variety of shapes and are referred to as *remnants* of the supernovae.

Every few centuries there is a supernova close enough to be seen with the naked eye, and some of these have been spectacular. On 1 May 1006, a new star appeared in the constellation Lupus and, within a matter of days, became the brightest star observed in all of recorded history. According to records kept by Chinese and Arabic scholars at that time, this star seemed ‘glittering in aspect, and dazzling to the eyes’. ‘The sky was shining because of its light’. ‘Its form was like the half Moon, with Pointed rays shining so brightly that one could see things clearly’. This nearby supernova (a very bright ‘new star’) was awe inspiring. It was probably visible for 3 months during daylight, and only after 3 years did it fade below naked-eye visibility at night. It captured the interest of more people than just the professional astronomers who, at that time, considered this a portent of natural disaster and political change.

At least seven times in the last 2000 years, such explosions have occurred in our Galaxy and near enough to Earth to appear as bright new stars. Clark and Stephenson (1977) review the historical records and the evidence for current identifications. The SNe are listed in Table 8.1. The last one was discovered on 9 October 1604, a few years before the first astronomical use of the telescope by Galileo, so none of the progenitor stars of these nearby supernovae were identified.

On 23 February 1987, to the great joy of modern astronomers, a supernova appeared in the Large Magellanic Cloud (LMC) – the first naked-eye supernova in 383 years. Although not as bright as the historical supernovae, SN 1987A was near enough to be able to identify the progenitor star: a B3 supergiant (a bright blue star). This was the very first time that past observations were available which revealed the appearance of the star before the explosion. Equally exciting was the detection of a burst of neutrinos marking the exact time of the event; a confirmation of theories about processes occurring deep in the core of a collapsing star, a region hidden from conventional observation.

Aside from the LMC supernovae, astronomers, routinely using telescopes having 10 000 times the light-gathering power of Galileo’s instrument, now find hundreds of supernovae a year in extragalactic nebulae. Figure 8.1 shows such a discovery. Although this was one of the brighter supernovae, it was not bright enough to be visible to the naked eye. Almost all extragalactic supernovae are so distant that it is impossible to locate the stars which explode to make them. Only a few,

Table 8.1 Historical and recent supernovae.

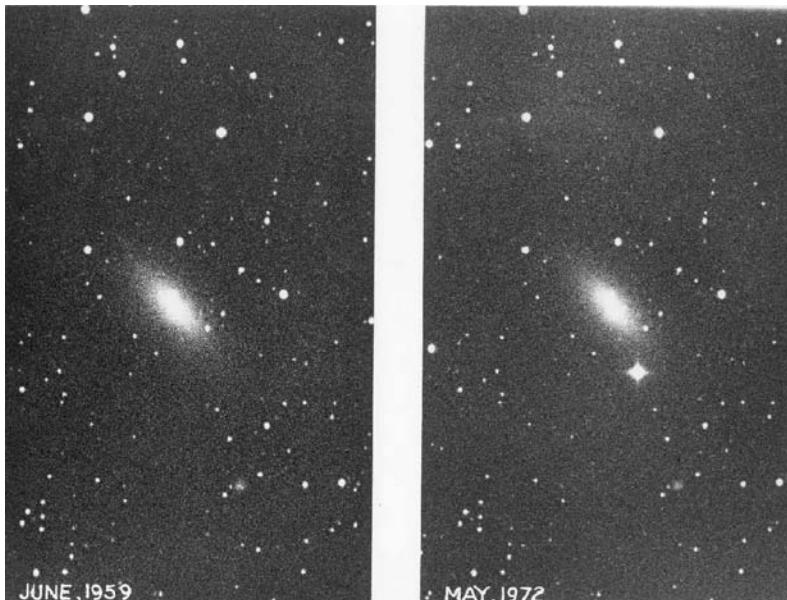
Date A.D.	Observers	Optical mag.	Type	Location	Remnant	Radio character
185	Chinese	-8	I?	Centaurus	RCW 86	Shell
1006	Many	-9.5	Ia	Lupus	SNR 1006	Shell
1054	Many	-4	II?	Taurus	Crab Nebula	Bright amorphous
1181	Chinese	0	II?	Cassiopeia	3C 58	Faint amorphous
1572	Tycho Brahe	-4	Ia	Cassiopeia	Tycho SNR	Shell
1604	Johannes Kepler	-3	Ia	Ophiuchus	Kepler SNR	Shell
≈1667	Nobody	?	II?	Cassiopeia	Cas A	Bright shell
≈1868	Nobody	?	Ia?	Sagittarius	G1.9+0.3	Shell
1987	Everybody	3	Ib	Dorado	SN 1987A	To be determined

including the progenitor star of SN 1987A, have been identified. The nature of most presupernova stars and the explosion mechanism must be deduced from theory and the study of events after the explosion. Such information includes optical spectra and light curves of the supernovae themselves and the structure of remnants that are typically hundreds or thousands of years old. These observations are compared with models and with theoretical calculations that tax the capability of the largest computers.

Supernova explosions are the most energetic stellar events known. The energy deposited in the debris and immediate surroundings is typically 10^{51} ergs. This is as much energy as the Sun

will radiate at all wavelengths in its 10-billion-year lifetime. Although the optical energy radiated during the first year is only 1 per cent of this, the brighter historical supernovae at their peak were visible during the day, and all could be seen at night for months or even years. Most of the energy is initially energy of motion. The outer layers of the star are thrown into space with initial velocity typically $10\,000\text{--}15\,000\text{ km s}^{-1}$. As the remnant expands, this energy of motion heats both the ejecta and the surrounding material to temperatures well above a million degrees, and X-rays are radiated profusely. Supernovae and their remnants were the only class of object predicted to be X-ray emitters *before* cosmic X-ray

Fig 8.1 Discovery of an extragalactic supernova: supernova 1972E in NGC 5253, a nearby elliptical galaxy in the Centaurus group. The supernova reached a peak brightness of eighth magnitude; only once every 30 years is a supernova of this brightness recorded. This galaxy was also the site of an earlier bright supernova in 1895 (photo from Hale Observatories).



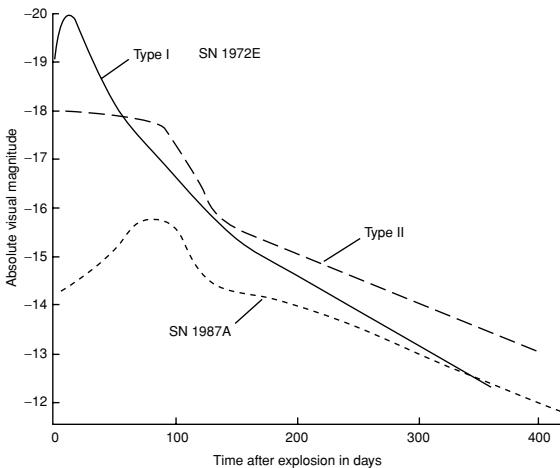


Fig 8.2 Light curves from three different supernovae. SN 1972E (solid curve) was typical of most type Ia SNe. The light curves are initially bright and fade exponentially at later times. Type II SNe (dashed curves) show more individual variation.

sources were discovered in 1962. Indeed, the Crab Nebula, a famous supernova remnant, was the first X-ray source to be identified with a previously known object.

To understand observations of the remnants, it helps to consider the late stages of stellar evolution. What leads up to these cataclysmic events? It turns out that two processes can lead to an explosion, each imprinting a signature on the optical supernova and upon the remnant.

8.2 | The supernova explosion

8.2.1 Two basic types

The two main types of SN explosions are known as type Ia and type II, with the distinction based on their observed optical properties, principally their light curves and spectra, as described in Box 8.1 and as illustrated in Fig. 8.2. The optical spectra of type Ia SNe show a variety of lines. Most are difficult to identify and all are heavily Doppler broadened because of the high initial velocities. There is a marked absence of hydrogen. Many type Ia events have similar light curves, suggesting a common type of progenitor star and explosion mechanism. The intensity rises quickly to maximum and reaches a luminosity of more than

$10^9 L_\odot$ in about 2 weeks. An initial rapid decay is followed by a long, slow decline in brightness. The luminosity falls exponentially with a characteristic time of about 55 days until the supernova fades to invisibility.

The spectra of type II SNe are dominated by broad emission lines of hydrogen, suggesting that the explosions occurred in young stars which had hydrogen-rich outer envelopes. Type II SNe rise more slowly to a maximum and are generally 2 magnitudes less luminous than are type Ia. The maximum is broader and the decays are on less regular timescales. A great deal of individuality characterises the decay of type II supernovae, implying a marked range of progenitor stars or explosion characteristics.

8.2.2 Where they are found

The distribution of type II events observed in distant galaxies supports this classification scheme. Type II outbursts occur in the arms of spiral galaxies, regions known to contain bright young (massive) stars and dense clouds of gas and dust characteristic of recent star formation. They rarely occur in elliptical galaxies where the stellar population is older. In contrast, the rarer type Ia events occur in all kinds of galaxies and show no preference for spiral arms. This suggests that the progenitor stars of type Ia supernovae are billions of years old and consequently not very massive. The type Ia light curves are remarkably similar, so much so that, knowing the shape, the absolute brightness of the SN at maximum can be accurately calculated. This has been used to determine distances of distant SN and is leading to new ideas about the expansion of the Universe (Perlmutter & Schmidt, 2003). The type II SNe show more variety. Many are classified as peculiar. Indeed, there is a school of thought that considers the type II phenomena to be a continuous gradation of characteristics rather than exhibiting discrete classes.

8.3 | Evolution to the explosion

Regardless of whether the supernova is of type Ia or type II, there is agreement that the cause is a catastrophic release of energy at the end of the life of a star. A star generates internal energy

Box 8.1 | Naming and classification of supernovae

The first supernova (SN) discovered in the year 2004 was SN2004A, the second was SN2004B, the 130th was SN2004dz (lowercase if two letters) and so on. This scheme indicates order of discovery, not order of time of explosion (as viewed from Earth). Supernovae (SNe) are classified on the basis of optical spectra and light curves (the change of brightness with time). At first, two broad categories were used, based on absence (type I) or presence (type II) of H. These corresponded to the two recognised mechanisms for the explosion of stars: (1) a $1.44 M_{\odot}$ white dwarf driven past the Chandrasekhar limit and (2) core collapse of a massive early-type star. It is believed that white dwarf progenitors are all similar and that all white dwarf explosions result in a type Ia SN. The optical spectra contain no H, but Si is prominent.

On the other hand, the core-collapse SN progenitors have initial masses in the range $8-25 M_{\odot}$ and comprise several types of stars. Consequently,

the core-collapse SN have a variety of characteristics and are given a variety of classifications. Some progenitors, such as Wolf-Rayet stars, have lost all their H. Resulting SN are type Ib – the spectra show no H, but He is present – and type Ic – the spectra show neither H nor He. Some core-collapse progenitors retain some of their H. These produce type IIL SNe – the spectra show H early but look like Ib or Ic later. Red supergiant progenitors have massive H envelopes and produce the following type II-P, strong H in the spectra and the light curves fall to a plateau soon after maximum; type IIn, strong H in the spectra and narrow emission lines; and type IIL, strong H in the spectra and the light curves fall linearly with time. The first SN progenitor to be identified was that of SN1987A, a blue supergiant. The spectrum was classified as ‘peculiar’, type IIPec – strong H in its spectrum but not a good fit to any of the preceding types.

by thermonuclear reactions. A high internal pressure and temperature support the outer layers and resist the star’s tendency to collapse under its own weight. At first hydrogen is fused in the core to form helium, then the helium is fused to form carbon and oxygen, and if the star is massive enough, fusion continues until the core is mostly iron. The nuclear binding energy of iron is greater than that of any other element, so further energy release through fusion is impossible.

Throughout this process, gravity has inexorably compressed the core of the star, creating the higher densities and temperatures that enable thermonuclear burning of the heavier elements. The internal energy generated eventually reaches the surface and escapes as radiation. When the fuel is exhausted, collapse continues until halted by other forces. A star the size of the Sun is stabilised before thermonuclear burning is complete by degenerate electron pressure. At this point the core consists of carbon and oxygen and is a million times denser than ordinary matter. The material has been compressed until all empty space has been squeezed out of the atoms.

At this point the electrons resist being crowded closer together, and collapse is halted. The surface of the star is now ‘white hot’ and it is called a white dwarf – a star with the mass of the Sun and the size of the Earth. A star with mass less than $1.44 M_{\odot}$, the ‘Chandrasekhar limit’, cannot evolve beyond this stage.

As the natural end point in the evolution of solar-type stars, white dwarfs are extremely common in our Galaxy. Examples abound, of which Sirius B, only 1.3 pc distant, is the closest and probably the best known.

8.3.1 Type Ia explosions: White dwarf ignition in a binary system

The similarity of optical characteristics of most type Ia supernovae implies similar progenitors. White dwarfs, at the Chandrasekhar limit, would provide a constant starting point. Such stars, however, are stable, and it is necessary to drive the star over the limit to trigger an explosion, probably by adding more mass. Thus type Ia progenitors could be white dwarfs in binary systems undergoing mass transfer. Material is added to

the white dwarf until the mass is high enough for gravity to overcome the resistance of electron degeneracy pressure. The resulting collapse raises the temperature until the carbon and oxygen in the core start to fuse. This creates an explosive wave, a deflagration (an explosion which proceeds without initiation by a shock wave), which propagates through the core in seconds. The resulting nuclear fusion reactions create a solar mass of radioactive ^{56}Ni and release 10^{52} ergs of energy, completely disrupting the star. The delayed release of energy via radioactive decay from the ^{56}Ni causes the expanding debris to glow and explains the slow decay characteristic of type Ia light curves.

This model thus accounts for the basic properties of type Ia supernovae. A problem to its acceptance, however, is the difficulty of detection of the predicted large mass of diffuse iron in the central regions of the remnants.

Figure 8.3 shows the possible evolution of an interacting binary that will culminate in a type Ia supernova explosion. Two stars of moderate mass (e.g. $1 M_{\odot}$ and $3 M_{\odot}$) are formed at the same time into a binary system (half of all stars are in binaries) and begin their main sequence evolution normally. The time in which a star completes its hydrogen-burning phase on the main sequence is inversely proportional to the square of the star's mass, so the more massive star evolves first. Once the central hydrogen has been exhausted, the helium core shrinks until the density is sufficiently high for helium burning to begin and provide the energy required to support the star against the force of gravity. The higher temperature, however, causes the outer layers of the star to expand. This material, instead of forming the usual red giant envelope, is either captured by the less massive star or flung into space. After this first stage of mass transfer is complete, the second star is left as the more massive star, and the first has become a carbon-oxygen white dwarf. The second star now evolves rapidly and starts to enter its own red giant phase, at which time it transfers matter back onto the white dwarf. The mass of the white dwarf increases until it passes the Chandrasekhar limit. At this point gravity initiates the collapse which leads to the nuclear detonation of the interior of the white dwarf.

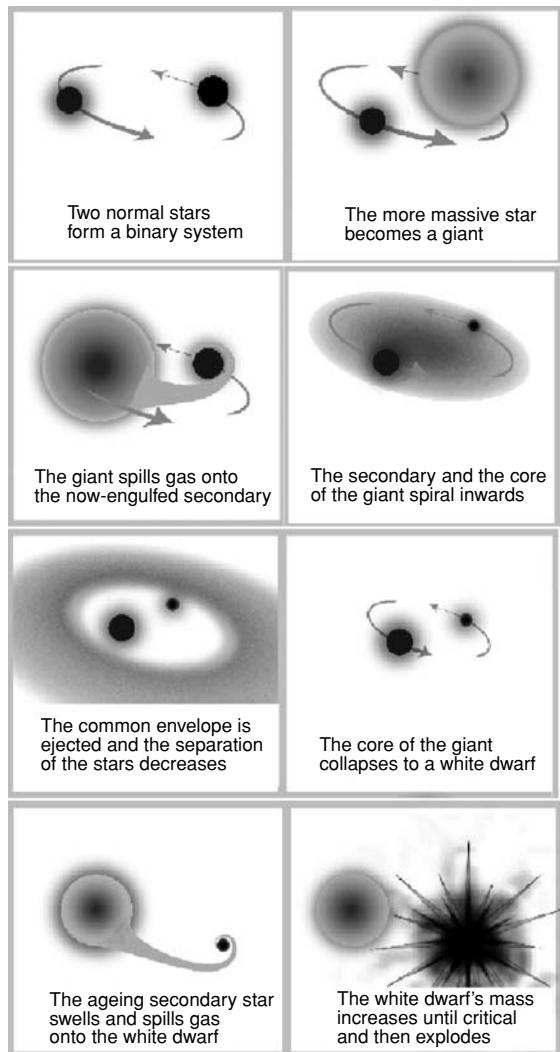


Fig 8.3 A way to make type Ia supernovae. The binary system evolves to produce a white dwarf and to then add mass until it is driven past the Chandrasekhar limit (adapted from Wikipedia, 'Supernova').

8.3.2 Type II explosions: Gravitational collapse of the stellar core

The evolution of a star starting its life with more than $\approx 10 M_{\odot}$ is different. It consumes nuclear fuel rapidly. A typical optical luminosity is 10^4 times that of the Sun or more. When its resources are exhausted, the end is marked by a great catastrophe: a type II supernova explosion. Gravity has been the dominant force in the life of the star,

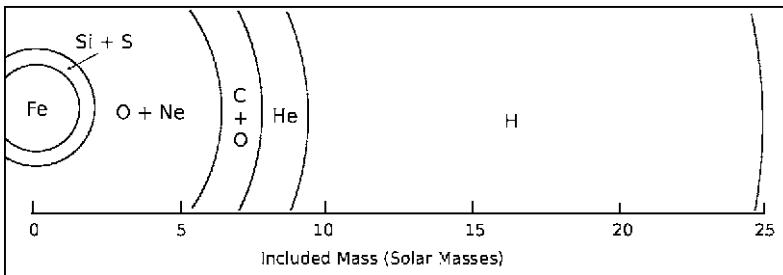


Fig 8.4 Elemental composition of a massive star towards the end of its evolution. The products of thermonuclear burning are in concentric layers surrounding an iron core. This diagram is not drawn on a linear scale but according to the mass included in each layer. Thus the Fe core contains $1.4 M_{\odot}$. Density falls rapidly with increasing radius. It is 10^9 gm cm^{-3} at the centre, 10^4 in the O/Ne/Mg layer and drops rapidly from 10^3 to 10^{-6} in the He layer. Density continues to fall in the outer H shell which occupies almost all the volume of the star (adapted from Brown & Bethe, 1985).

and the energy released in the explosion is gravitational also. At the beginning, hydrogen in the core is fused to helium. When this is depleted, the star contracts until the temperature is high enough to fuse helium into carbon and oxygen. At the same time the pressure increases in the layer of hydrogen surrounding the core. This results in initiation of nuclear burning in this outer layer, creating a shell of helium surrounding the core of heavier elements. At the end, after a life of perhaps only 10 million years or less, the star is stratified into onionlike layers (Fig. 8.4). The central core is now iron, the end point of the fusion process. This is surrounded by shells of silicon and sulphur, neon and oxygen, carbon, helium and an outermost shell of hydrogen. The density in the core is that of a white dwarf whereas the outer layer is very tenuous. Most stars at this phase are probably red giants or supergiants surrounded with glowing shells of hydrogen as large as the inner part of our Solar System. We cannot see through these hydrogen shells to observe the inner part of the stars where the final stages of evolution rapidly occur.

The iron core can no longer generate nuclear energy. It is stabilised by electron pressure and is at the Chandrasekhar limit. The mass of the core, however, is continuously increasing as the adjacent layer of silicon is fused into more iron. The core is compressed more and the internal

temperature increases. This causes some of the iron to decompose into lighter nuclei. As a result, energy is absorbed, reducing the pressure and causing the core to shrink. Free protons are created which combine with electrons to make neutrons and neutrinos. Some of the electrons supporting the core thus disappear, causing further pressure drop. The process now runs away, and gravity overwhelms the electron pressure. Within milliseconds the core collapses until nuclear density is reached: a solar mass within a radius of 10 km. This process is difficult to believe. The mass is slightly larger than that of the Sun. The gravitational field is enormous and the rapidity of the collapse surprising (only a few seconds). Yet on 23 February 1987 the neutrinos created in such a collapse were actually observed as the first signal from SN 1987A.

The energy explosively released by the in-falling matter creates a shock wave. This shock propagates outward through the still in-falling outer layers until it reaches the outermost part of the star. The shock ejects this last material outwards. We observe light from the expanding hot material as a type II supernova. A rapidly cooling hot neutron star is left at the centre of the explosion, and if this remnant of the core is spinning rapidly, it is later observable as a pulsar at the centre of the remnant. Total energy generated by the collapse and subsequent nuclear changes is in excess of 10^{53} ergs. This is largely carried off by neutrinos, emitted as electrons and protons combine to form neutrons. The neutrinos, because of their low reaction cross section, escape easily from the collapsing star but are therefore also extremely difficult to detect. The light from the supernova itself carries only 10^{49} ergs of energy. The kinetic energy of the expanding debris is typically 10^{51} ergs. Rotational energy of the neutron star can also be appreciable; if spinning at 30 cycles s^{-1} , this is 2×10^{49} ergs.

8.4 X-rays from supernovae

There are several mechanisms that can generate X-rays early in the life of a supernova. At the time of the event, there may be prompt X-rays connected with the breakout of the shock from the stellar surface or from the not-well-understood γ -ray burst (GRB) process (see Chapter 17). There could be X-rays from a newly formed neutron star (in about a year the expanding debris cloud could become transparent enough to see through to the centre). Gamma-rays from short-lived radioactive nuclei in the debris can be Compton scattered into the X-ray waveband. And finally, the ejected outer layer ploughing into surrounding material can be heated to temperatures high enough to emit X-rays. Indeed, it is this mechanism that accounts for the emission from almost all the older remnants. Prompt X-rays have recently been detected from the shock breakout. X-rays have also been observed at later times (within days to years of the explosion) from about a dozen SNe and are well explained by the ejecta-circumstellar material interaction.

8.4.1 Shock breakout

There has been one observation of a remarkably brief (<7 min) X-ray flash which marked the beginning of SN 2008D in the normal spiral galaxy NGC 2770 on 9 January 2008. This was observed accidentally by Soderberg *et al.* (2008) in attempting SWIFT follow-up observations of SN 2007uy, which had erupted in the same galaxy only weeks before. Figure 8.5 shows the time history of the X-ray flash. It was extraordinarily intense, with peak luminosity of $\sim 6 \times 10^{43}$ ergs s^{-1} (at the galaxy's distance of 27 Mpc). It faded rapidly, leaving a UV and then optical afterglow. This was subsequently swamped (within a week) by the visual emission from a type Ibc supernova explosion.

The X-ray burst was the theoretically predicted, but never seen, signature of shock breakout from the surface of the star (Colgate 1974). Upon collapse of the core, a shock is generated which propagates outward from the newly created neutron star. When the shock leaves the optically thick material below the surface, it propagates

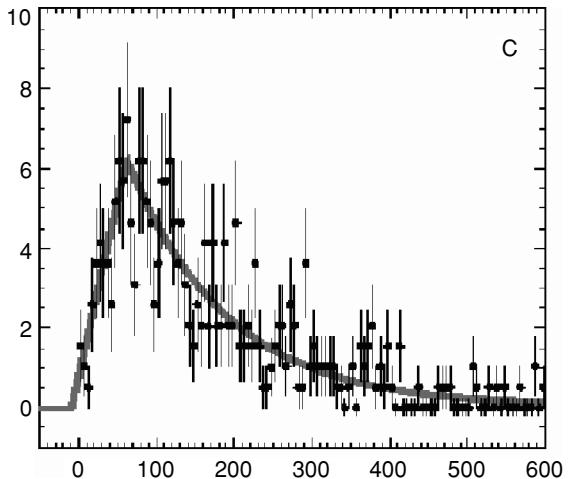


Fig 8.5 A burst of X-rays observed at the start of SN 2008D. Count rate of the SWIFT detector in the energy range 0.3–10 keV is plotted as a function of time since the start of the burst. The duration of the event was 7 min. The solid curve is a fast-rise exponential-decay model (from Soderberg *et al.*, 2008).

through material ejected as a wind just hours before the explosion. This material is optically thin and strongly heated. The resultant pulse of X-rays contains information about the density and distribution of the wind. In this case, data indicate a Wolf-Rayet progenitor star.

Total X-ray energy of this burst was $\sim 2 \times 10^{46}$ erg. This is three orders of magnitude below the energy of a weak γ -ray burst (Chapter 17). It is a different phenomenon and is called an X-ray outburst (XRO 080109). Together with the pulse of neutrinos emitted (only ever seen for SN1987A), this X-ray pulse defines the start of the supernova event precisely. The XRO from SN 2008D occurred 9 January 2008 at 1332:40 \pm 20 s UT.

8.4.2 SN 1987A

SN 1987A was promptly classified as type II. The optical spectrum clearly showed the presence of hydrogen. After some initial confusion, the progenitor star was identified as the 12th-magnitude star Sk-69 202, a B3 supergiant and well documented in previous photographic plates. This was a young star of about $20 M_{\odot}$ but not the red supergiant progenitor expected. There is no doubt that the event originated in gravitational collapse of

the core of the star. This was established by the simultaneous detection of a burst of neutrinos by two Earth-based detectors. These detectors, it is worth noting, were not made for astronomical research but were built to search for extremely rare proton decays predicted by some theories of fundamental particles. They were 10-m cubes of pure water surrounded by photomultiplier tubes and buried deep in the Earth to reduce the background noise from cosmic rays. The detector at Kamioka, Japan, recorded 11 neutrino-induced events in a span of 13 s. The Irvine-Michigan-Brookhaven detector at Cleveland recorded eight events in 6 s, a ‘burst’ in a system having a background of just one event every 5 days. Thus we know that collapse occurred at 0735 UT 23 February 1987, 18 hours before Ian Sheldon from the University of Toronto started the exposure of the discovery photograph. The detection of these neutrinos was a triumph of both theory and observational science. The measurements yielded the neutrino flux which gave the energy release in the collapse, 3×10^{53} ergs, exactly the gravitational binding energy of a neutron star. The fact that the neutrinos did not arrive simultaneously but over an interval of several seconds, implies many interactions in the collapsing core before escape. The timing of the observed events was used to study the motion of material during formation of the neutron star.

8.4.2.1 Fainter than expected?

The optical light curve observed was unusual but was explained by the nature of the progenitor star. A prediscovery photograph showed appreciable brightening only 3 hours after core collapse. The star brightened rapidly, reaching a maximum of magnitude 4.5 in 4 days. After a small decline in brightness, there was a slow rise to maximum light of magnitude 3.0, 70 days after discovery. A ‘typical’ type II supernova in the LMC would have reached magnitude 0 or 1 at maximum brightness. The fact that the progenitor lacked the extensive envelope of a red supergiant accounts for the very rapid initial rise and for the maximum brightness of a factor of 10 less than expected. Light from the supernova then faded exponentially, powered by the radioactive decay of ^{56}Ni created in the explosion. ^{56}Ni decays with a half-life of 7 days to ^{56}Co ,

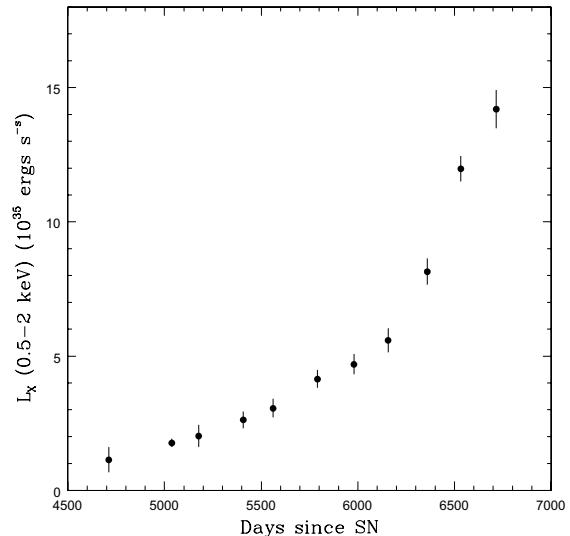


Fig 8.6 X-ray emission from SN 1987A measured by Chandra from January 2000 to July 2005, showing how the X-ray flux increases as the shock wave sweeps up and heats the surrounding medium (from Park *et al.*, 2006).

which in turn decays with a half-life of 77 days into stable ^{56}Fe . There was speculation about a possible contribution from a rapidly spinning neutron star at the centre, but this has not yet been seen.

8.4.2.2 X-ray emission

X-rays at energies 6–16 keV were detected by Ginga in the interval 130–340 days after stellar collapse (Dotani *et al.*, 1987; Inove *et al.*, 1991). X-rays with a luminosity $\approx 10^{37}$ ergs s^{-1} were also detected in the range 10–100 keV with a high-energy detector on the Soviet Mir-Kvant spacecraft (Sunyaev *et al.*, 1988). At this time the expanding shell was still so thick that X-rays from within could not escape. Hence those detected were originally γ rays from the decay of ^{56}Ni and ^{56}Co , downgraded in energy to X-rays by collisions with electrons in the expanding shell.

Monitoring started again with the launch of ROSAT. Weak, soft X-rays were detected at an age of 4 years. At this time the expanding debris was transparent to radiated X-rays and $L_X \approx 10^{34}$ ergs s^{-1} . By 8 years, there had been a steady increase to 2×10^{34} ergs s^{-1} . Chandra observations started after 13 years and the increasing emission over the next 5 years is shown in Fig. 8.6. At age 18 years (July 2005) L_X had risen to $\approx 1.5 \times 10^{36}$ ergs s^{-1} .

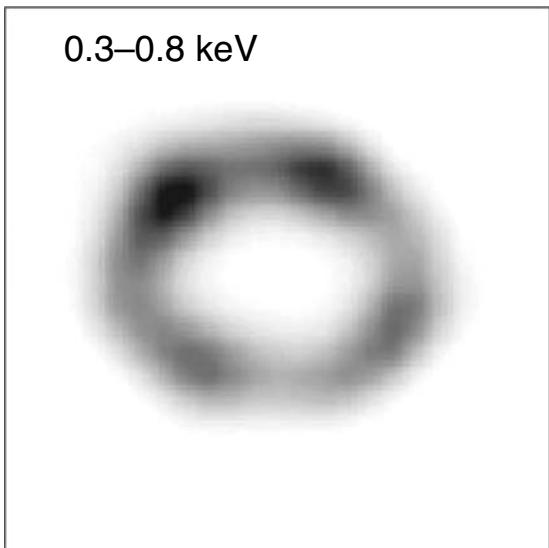


Fig 8.7 X-ray emitting region of SN 1987A observed July 2005. The ACIS image from a 44-ks exposure has been de-convolved using the instrument response to a point source. The observed ring was 1.6" in diameter (from Park *et al.*, 2006).

Figure 8.7 shows the Chandra image of SN 1987A. Emission is from a ring-shaped region of diameter 1.6". This ring is a structure formed by the wind history of the precursor star. A slow wind followed by a fast wind formed a relatively dense shell around the star. This shell is apparently densest in the equatorial region.

X-rays come from shock-heated material ahead of the expanding ejecta. When the shock reaches this shell, X-ray emission will greatly increase. The rapid rise in emission starting at 6000 days indicates that this increase has started. Figure 8.8 shows the spectrum which is interpreted as emission from plasma at two different temperatures. The lower temperature is from material of the inner ring heated by a shock which has been slowed by dense material. The higher-temperature emission is from a fast shock in less-dense material around the ring. Emission lines from several elements can give information on the temperature and velocity of the heated material (see Fig. 8.9).

Thus SN1987A was a bright X-ray source for its first year, with the energy coming from decay of radioactive elements in the expanding debris. This emission faded, and weak emission from the collision of debris with circumstellar material subsequently appeared. These shock-heated X-rays have since increased in luminosity by a factor of 150 and are showing the distribution of circumstellar material. SN1987A is now bright enough so that high-resolution spectra can be obtained in only 100-ks observing time. Temperature, composition and velocity of material are being derived directly from X-ray observations.

8.4.3 Supernova 1993

On 18 March 1993 a supernova was discovered in M81, a well-studied, nearby spiral galaxy. Six days later, X-ray emission was easily detected by

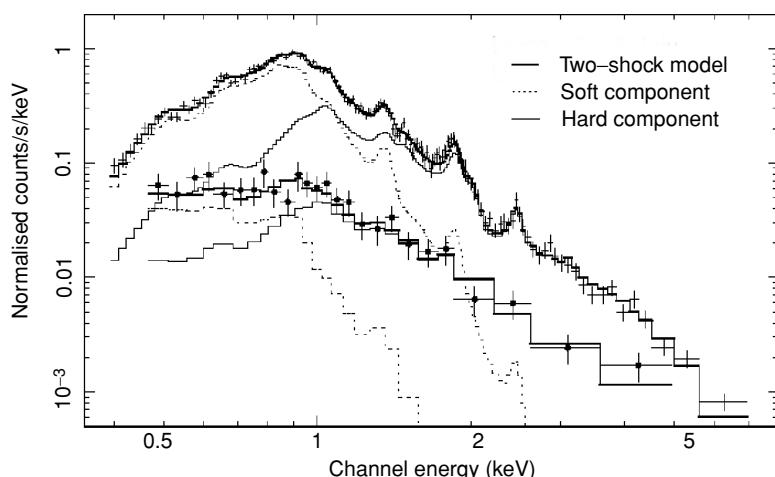
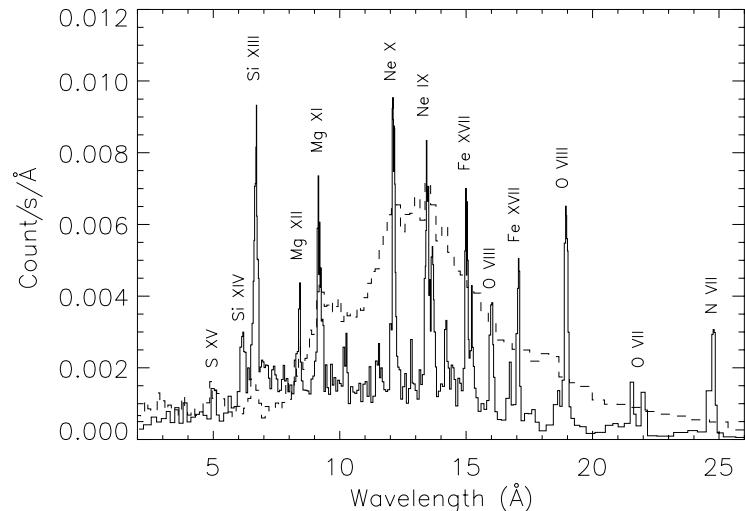


Fig 8.8 X-ray spectra of SN 1987A from the Chandra ACIS-S detector measured over a 5-year interval from (bottom) January 2000 to (top) July 2005. Solid lines are best-fit spectra calculated using a two-shock model. Dashed lines show contributions from low- and high-temperature regions (from Park *et al.*, 2006).

Fig 8.9 High-resolution X-ray spectrum of SN 1987A obtained with the Chandra Low Energy Emission Grating (LETG)/ACIS spectrometer. This is the sum of a series of observations taken August 2004 with total exposure of 289 ks. The dashed line shows the zero-order ACIS pulse height spectrum. Prominent lines in the dispersed spectrum are labelled. To quote the authors, ‘This figure demonstrates the enormous scientific advantage of the dispersed spectrum over the pulse-height spectrum’. (from Zhekov *et al.*, 2005).



ROSAT, and the SN was quite bright. The X-ray luminosity was 2×10^{39} ergs s $^{-1}$. The SN was monitored for 4 years by ROSAT and later observed by ASCA, Chandra and XMM with results shown in Fig. 8.10. SN 1993J began as a type II, but after a few months the hydrogen emission faded and spectra indicated type Ib. The precursor star was identified as a K0I supergiant with probable mass when formed of $\approx 15 M_{\odot}$; at the time of the explosion, it had probably already shed most of its outer tenuous atmosphere of hydrogen. Because much of the mass loss was through a stellar wind, the precursor was therefore surrounded by material that had been blown away in the wind before the explosion.

During the first year this SN was much brighter in X-rays than SN 1987A. Even at early times, the X-ray emission from SN 1993J was probably from ejecta ploughing into nearby circumstellar material. The distribution of this surrounding material can be deduced from the time history of the subsequent X-ray emission. In addition to the X-rays, SN 1993J was one of the brightest radio-emitting supernovae yet observed, and the radio luminosity also gives information about the distribution of circumstellar material.

X-rays come from a shell-like region behind a shock propagating into circumstellar material. If the circumstellar material were uniform, the emission would increase with time, indicating the increase in emitting volume as the shock moves outwards. In this case the observed decrease in

X-ray emission with time shows that the density of circumstellar material *decreases* with distance from the explosion site. In contrast to this, the radio emission at first increased with time, showing that the emitting region was initially almost opaque to radio waves. The expanding shell became transparent to radio waves after 100 days, in contrast to only a few days required for X-ray transparency. The X-ray light curve has therefore been used to infer the stellar wind history of the progenitor in the years leading up to the explosion (Immler *et al.*, 2001). Assuming a steady red-giant wind with velocity ≈ 10 km s $^{-1}$, the mass loss rate must have decreased constantly during the last 10^4 years of evolution in the manner shown in Fig. 8.10. This behaviour indicates that the star was in the process of changing from a red to a blue supergiant at the time of the explosion.

Thus the observed X-ray emission can be used to derive information concerning the progenitor. The circumstellar environments of the preceding two examples are quite different. We are just starting to learn about SN 1987A.

8.5 Evolution of supernova remnants

Most galactic remnants were discovered as spatially extended sources in radio surveys. They were distinguished from HII regions (clouds of ionised

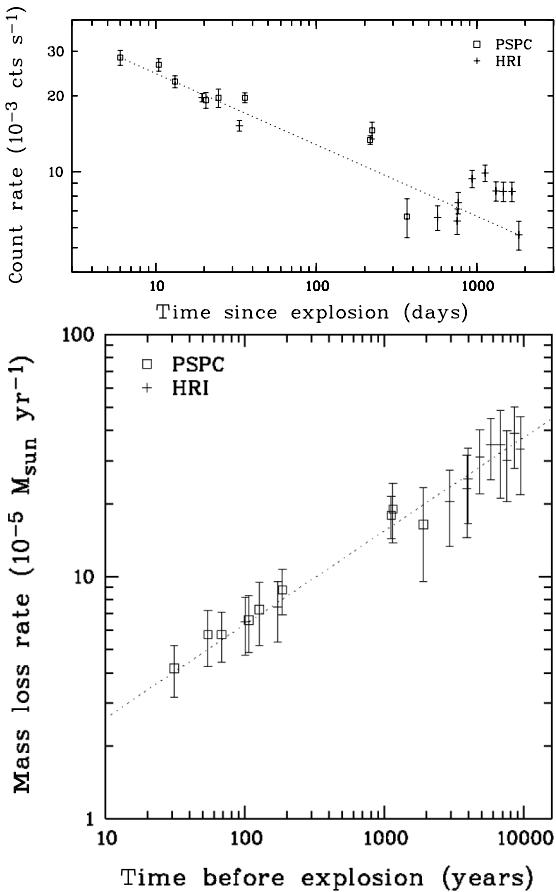


Fig 8.10 (top) X-ray emission of SN 1993J as measured by ROSAT over a 5-year period. (bottom) Mass loss of the precursor star to SN 1993J as deduced from the ROSAT light curve (from Immler et al., 2001).

hydrogen usually associated with luminous hot stars) by the relatively flat shape of the radio spectra, a shape indicating that the radio emission mechanism in SNRs is *not* a thermal process. Many remnants carry the names assigned after the radio discovery. Cas A and Tau A (better known as the Crab Nebula) are the brightest radio sources in these two constellations, and Vela XYZ is formed of three radio-bright regions. Catalogues of such non-thermal extended radio sources now contain about 200 objects in our Galaxy (Green, 2006) and about 50 in the Magellanic Clouds. Most of these sources are somewhat circular in shape and are limb brightened, as expected of emission from a large, hollow shell which is transparent to its

own radiation. This is the simplest shape expected for the expanding debris from a supernova explosion.

8.5.1 Phase I: Free expansion

The very simplest model assumes material ejected equally in all directions from a progenitor star embedded in a uniform medium. The evolution is shown schematically in Fig. 8.11, and it is customary to consider three phases in the expansion of the remnant. The shell of ejected material first expands rapidly and sweeps up the surrounding medium like a snow plough, leaving a low-density region behind in the interior. The density of material in the rapidly expanding shell decreases with time and soon becomes low enough for the shell to be transparent or ‘thin’. During this phase the mass of swept-up material is negligible compared to the mass of the ejecta, and the expansion proceeds at a uniform velocity. This ‘free expansion’ is the first phase in the life of a supernova remnant and has the following typical properties: if $1 M_{\odot}$ is ejected and the density of the surrounding medium is $0.3 \text{ atoms cm}^{-3}$, this phase will last until the radius is about 3 pc, when the swept-up mass becomes equal to that of the ejecta. If the initial velocity is 15000 km s^{-1} , then the age of the remnant at this time will be 200 years.

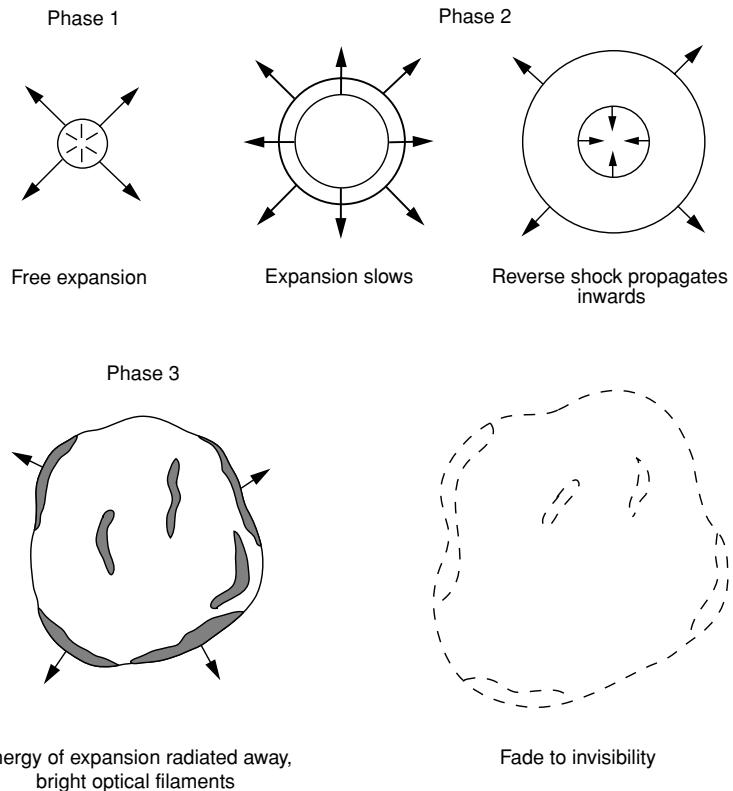
8.5.2 Phase 2: The blast wave or Sedov-Taylor phase

As time passes, an increasing amount of material is swept up by the shock and the remnant enters the second phase, the Sedov-Taylor or *blast wave* expansion. The mass of swept-up material is now large compared to the mass of the ejecta. The energy radiated by the material in the shell is still small compared with its internal energy, so the expansion is adiabatic. The rate of expansion is determined only by the initial energy deposited by the explosion E_0 and the density of the ISM n .

8.5.2.1 The reverse shock in the ejecta

If there were no material surrounding the SN, the stellar ejecta would continue to expand freely at $10000\text{--}20000 \text{ km s}^{-1}$. The interstellar gas, however, forms a barrier that is increasingly difficult for the expanding ejecta to push aside. As

Fig 8.11 Evolution of a supernova remnant. The young remnant first expands freely into the surrounding interstellar space. As the size increases, it incorporates an ever-increasing amount of interstellar material, radiating continuously, until it cools and is no longer distinguishable from the background.



Box 8.2 | The shock wave in circumstellar material

The speed of sound in a gas with temperature T and average molecular mass μ is $\gamma kT/\mu$. For an ideal monatomic gas, the ratio of specific heats at constant pressure and constant volume, γ is $5/3$. In the cold ISM the speed of sound is $1-10 \text{ km s}^{-1}$, which is very much less than the velocity of the stellar ejecta. A shock wave consequently forms at the leading edge of the ejecta and travels before it into the interstellar material. An atom or ion of interstellar hydrogen far from the star will suddenly find itself violently perturbed as the shock passes. The gas density just behind the shock is increased by a factor $(\gamma + 1)/(\gamma - 1) = 4$. Conservation of momentum at the shock front requires that for a shock of velocity v , material behind the shock is propelled outward, in the direction of motion of the shock, at a velocity $3v/4$. The temperature of material behind the shock is raised to $T = 3\mu v^2/16k$.

Behind an SNR shock the temperature becomes 10^7-10^8 K , causing the atoms to become largely ionised.

Within Earth's atmosphere, shock characteristics are well understood and shocks propagate via collisions of molecules (as demonstrated in Fig. 8.12). However, the ISM is more complicated. The density is a factor of 10^{21} less, and collisions are rare. Magnetic fields, however, are strong enough to moderate the process. Energy in a shock in the ISM is not transferred by collisions between ions and atoms. The energetic ions behind the shock transfer energy to material in front of the shock via the magnetic field. SNR shocks are thus called 'collisionless' shocks. Magnetic fields compressed by the shock account for the non-thermal radio emission observed from the shells of most SNRs, so their existence is not in doubt.

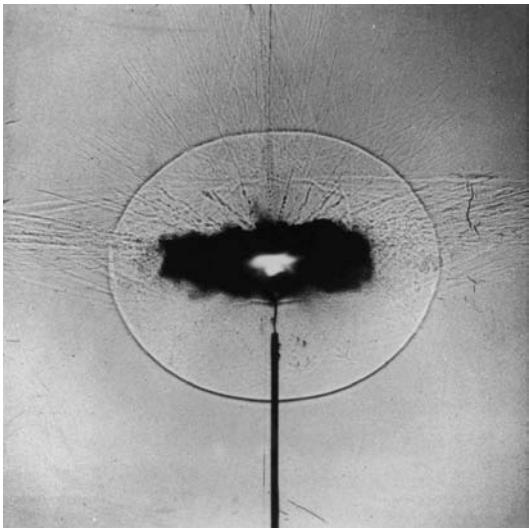


Fig 8.12 A shock wave formed by a small explosion in the atmosphere. A fraction of a second after the detonation of a dynamite cap, the gaseous products of the explosion are black and surrounded by an outgoing shock. Supersonic particles traveling ahead of the shock are also clearly seen. The same phenomena are visible in supernova remnants, even though the energy release is 10^{41} times greater (picture by Harold Edgerton).

the rapidly moving material ploughs into the surrounding medium, the interstellar gas is swept up and moves out with the ejecta. Two shock waves then form (Fig. 8.13). One propagates into the gas

ahead of the ejecta, and a second, the *reverse shock*, propagates back into the ejected gas. The boundary between circumstellar material and ejecta is called the *contact discontinuity*. As seen by an outside observer, both shocks initially travel outward, with the radii of the two shocks differing by about 25 per cent. After the swept-up mass exceeds that originally ejected, the reverse shock propagates inwards.

Only between the two shock waves is the material hot. Here the interstellar gas has been heated and compressed by the expanding ejecta. The ejecta have in turn been slowed and compressed by the pressure of the interstellar gas. At large distances the interstellar gas is cool and does not know about the expanding debris from the stellar explosion. In the central region the material is no longer hot and is freely expanding. The information that there will be resistance ahead has not yet been received. Only the shocked material is hot enough to emit X-rays, and this forms the bright shells observed in young remnants.

8.5.2.2 Rayleigh-Taylor instability and clumps

In nature the geometry is complicated by irregularities in both the surrounding medium and in the ejecta. The region around a SN can contain rather dense molecular clouds as well as fluctuations in density of interstellar gas. There may

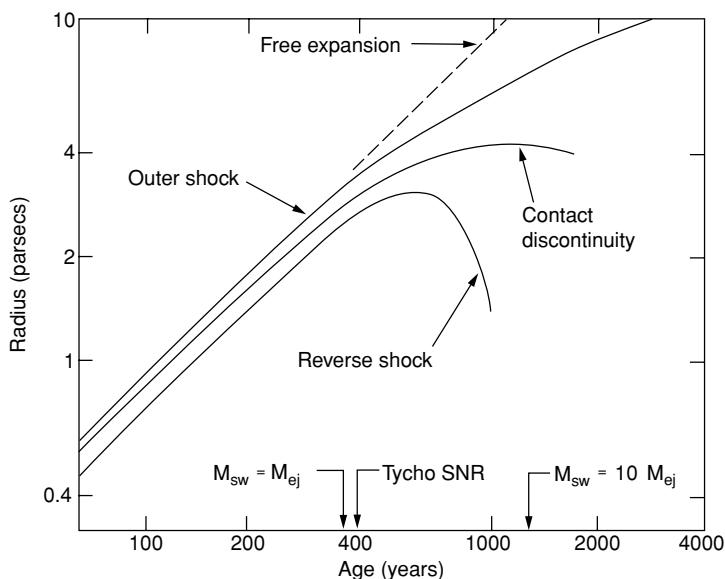


Fig 8.13 Position of the shock waves in a young remnant as it expands into a uniform surrounding medium. Numbers were picked to model Tycho's remnant: swept-up mass equals ejected mass after 370 years and after 1300 years, it is 10 times the ejected mass (based on the work of Chevalier, 1982).

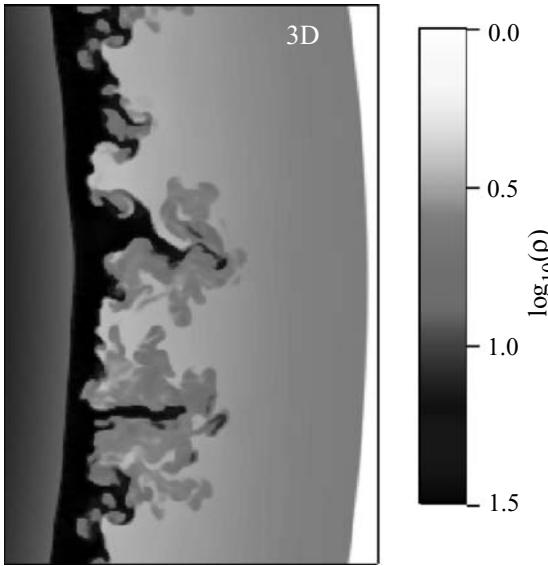


Fig 8.14 The R-T instability: a planar slice from a three-dimensional calculation of the shock structure in a supernova remnant. Material is travelling left to right. The scale shows density increasing going from white to black. The sharp grey-white boundary to the far right is the forward shock. The sharp grey-black boundary to the left is the reverse shock. The R-T instability has caused the contact discontinuity to form fingers and eddies and to mix with the hot gas behind the forward shock. Compare with the observation in Colour Plate 19 (from Blondin & Ellison, 2001).

also be structure produced by the varying stellar wind of the progenitor. Since thermal X-ray emission is proportional to the square of the density, small variations can produce large variations in luminosity of the hot gas. The structure of older remnants is therefore expected to reflect the properties of the circumstellar medium.

Even in remnants resulting from symmetric explosions, the ejected material is subject to the Rayleigh-Taylor (R-T) instability and is expected to break into clumps. This instability is commonly encountered on Earth as it occurs when a heavy fluid is supported by a lighter one (e.g. a large blob of water falling through air will break into small droplets). In an expanding remnant, the ejecta push against the shocked ISM at the contact discontinuity (CD). The R-T instability causes fingers of ejecta to form and to protrude ahead of the originally spherical CD. These fingers are themselves unstable, and there is turbulence which mixes

ejecta with shocked interstellar gas (as demonstrated in Fig. 8.14). In an adiabatic expansion, the turbulence is expected to extend halfway to the forward shock (FS). In the case where cosmic ray acceleration removes energy from behind the FS, the fingers can extend to and beyond the FS (Blondin & Ellison 2001). All this is applicable to the remnant of Tycho's SN, as seen in Colour Plate 19. Even young remnants appear clumpy and fragmented rather than projections of a simple spherical geometry.

Box 8.3 | The Sedov-Taylor model

The average remnant, between the ages of about 100 and 1000 years, sweeps up mass rapidly and cools as it expands; yet for some time, the loss of energy through radiation is small. During the adiabatic expansion, this model may be applied. It is appealing because it predicts a precise behaviour of the remnant with time, dependent only on E_0 and the density of the surrounding medium n . If the distance is known, the age, t , and E_0 can be derived from the measured X-ray flux and spectrum (which directly yield X-ray luminosity L_X and temperature T).

The radius R of the blast wave is

$$R = 14(E_0/n)^{1/5}t^{2/5}\text{pc},$$

where E_0 , n and t are in units of 10^{51} ergs, per centimeters cubed, and 10^4 years, respectively. The shock temperature is also given as

$$T = 1.0 \times 10^{10}(E_0/n)R^{-3}\text{K}.$$

L_X is related to the calculated (and tabulated) emissivity, $P(E, T)$ ergs cm³, of the hot gas. Both are dependent on temperature and X-ray photon energy E . The density distribution of gas in the shock also requires a dimensionless factor $q(T)$, which has a value of approximately 0.6:

$$L_X(E, T) = (16\pi/3)n_i n_e R^3 P(E, T)q(T) \text{ ergs s}^{-1}.$$

Thus, ignoring asymmetries and non-equilibrium conditions, this model can be used to derive SNR ages and energy release. This has been useful, although results are not exact (see Cox, 1972).

8.5.2.3 Phase 3: Dominated by radiation

As it expands, the remnant sweeps up cold interstellar material and becomes cooler as its mass increases. As the material behind the shock cools, the rate of radiation of energy increases. This somewhat paradoxical result arises because, after the temperature drops to $\sim 2 \times 10^5$ K, some electrons have recombined with the carbon and oxygen ions, and the gas is then able to radiate by the very efficient process of UV line emission. The remnant has now reached the third phase, the ‘radiative’ phase. During the 10^5 -year duration of this phase, most of the internal energy will be radiated. The shell will coast through interstellar space, becoming fainter and fainter until it is indistinguishable from the surrounding medium. This phase is sometimes called the ‘constant-momentum’ phase, even though momentum is, of course, conserved throughout the life of the remnant.

8.5.3 Behavior of ‘real’ SNRs

Generally, in different wavebands, the radiation is produced by different processes. All are stimulated by, and are most prominent, in the vicinity of the shock. X-rays are emitted most profusely during phases 1 and 2 from the hot material behind the shock. During phase 3 the bright optical filaments characteristic of older remnants form. Radio energy is radiated from the vicinity of the shock and from the cooling filaments and is usually detectable throughout the life of the remnant. Thermal X-rays are generated, as explained in Chapter 2, by fast electrons colliding with positive ions or by electron-ion recombination. The temperature of this plasma is 1 million degrees or higher. The optical emission comes from material with a temperature of $10\,000^\circ$ and consists of discrete optical and ultraviolet lines; all have wavelengths characteristic of the radiating atoms. The radio emission is synchrotron emission from high-energy electrons moving through a moderately strong magnetic field. Thus optical and X-ray emissions do not come from the same material. Remnants which exhibit both (which most do) contain material spanning a large temperature range and probably pervaded by high-energy radio-emitting electrons. Some remnants are also bright IR sources. Dust immersed in the hot gas

is heated until it radiates in the IR. Maps of the brighter remnants indeed show that the origin of most of the IR is within the X-ray-emitting shell. Total energy radiated in the IR is sometimes greater than in the X-ray band, and this will speed up the evolution of these remnants.

Of course, no remnants show the precise spherical shape assumed in the preceding discussion. Bright regions exist as shreds, patches and filaments reflecting sometimes large asymmetries in the distribution of ejected material or in the surrounding medium. Mathematical models use a simple geometry to derive a solution, but nature is more complicated.

8.6 | Young shell-like remnants

Chances of learning the nature of the progenitor star are best when the remnant is young. At this time its characteristics are determined by the nature of the star and the explosion. The observed material will be predominantly ejecta. A measure of the mass and composition of the remnant will then give the mass of the progenitor star and might reveal heavy elements made in the explosion. The energy released can be calculated using the observed mass of material and the velocity of expansion. Any asymmetries in spatial distribution or composition must be caused by the explosion itself. The youngest remnants known in our Galaxy are those listed in Table 8.1, but only SN 1987A and G1.9+0.3 are young enough to be expanding freely. The others are probably in a transition between phase 1 and phase 2, so the interaction with surrounding material, although not yet dominant, cannot be ignored. Recent X-ray observations of young remnants have clarified our view of the process.

8.6.1 The remnant of Tycho’s supernova (type Ia)

8.6.1.1 Discovery

The Danish astronomer Tycho Brahe first saw the supernova of 1572 on the night of 11 November. He was outside, after dinner, planning a night of observing after a stretch of bad weather. When he spotted the new star, which was close to maximum brightness, he could not believe his eyes.

Box 8.4 | Ionisation equilibrium of material in remnants

Shock waves in the shells of supernova remnants are collisionless. Energy is transmitted through the magnetic field to the surrounding material, which has been ionised by UV radiation from the supernova itself. To an observer riding on the outward-moving shock, the kinetic energy of incoming material resides in the positive ions which carry almost all the mass. This kinetic energy is thermalised by the shock, and immediately after passage of the shock, the energy is found in rapid motion of the positive ions which are still in a state (characterised by the number of bound electrons) corresponding to a much lower temperature than that characterising their motion. After a time, the free electrons, through collisions, share the energy of motion and come into thermal equilibrium with the positive ions. As more time passes, the now fast-moving electrons collide with and remove more electrons from the heavier positive ions, and eventually, the state of ionisation is increased to that appropriate to the electron temperature. Thus the three energy indicators of the gas, namely, positive ion velocity, electron velocity, and ion state, all reach the same temperature. The time required for this depends on

density and temperature and is typically hundreds of years.

Information about the state of the gas is obtained from the X-ray spectrum. The high-energy continuum shows the velocity of free electrons, and the emission line intensity and energy give the nature of the ions. Usually these are characterised by different temperatures. ‘Young’ remnants are not yet in equilibrium, and spectral analysis models take this into account with an additional parameter, the ‘ionisation timescale’ ($= n_e t$), the product of electron density and age of the plasma. The value for a young SNR is usually in the range $1 \times 10^{10} - 3 \times 10^{11} \text{ cm}^{-3} \text{ s}$, and when a plasma is close to equilibrium, $n_e t \approx 3 \times 10^{12} \text{ cm}^{-3} \text{ s}$. Such models are called non-equilibrium ionisation (NEI) models.

It may be helpful to note that in the literature, there are three ways of referring to ions. For example, an O nucleus with a single bound electron is called hydrogen-like, or O^{VIII}, or O⁷⁺. An O nucleus with two bound electrons is helium-like, or O^{VII}, or O⁶⁺. Using this nomenclature, in astronomy, it is customary to call hydrogen-like H, H_I and ionised H, H_{II}.

He was a seasoned observer, and in his experience the stars were fixed features in the arrangement of the heavens. This was also the belief of everyone else at the time. He asked others with him to verify that the new star was really there before going to his observatory and starting the first of many careful observations.

Tycho measured the brightness of the new star from maximum until it faded below naked-eye visibility 15 months later. He obtained the position by measuring the angular distance between the new star and the bright stars in Cassiopeia. His position is within 2 arcmin of the centre of the remnant we see with modern instruments. Tycho’s light curve is clearly that of a type Ia supernova.

The remnant of this supernova was discovered in the radio band 378 years after the supernova had faded to invisibility (Hanbury-Brown &

Hazard, 1952). Extremely faint optical wisps at this location were found on a photograph taken in 1949 (van den Bergh, 1971). The X-rays were first recorded with a rocket-borne detector in 1965 (Friedman *et al.*, 1967), and the X-ray source was identified as the remnant in 1968 (Gorenstein *et al.*, 1970). Detailed radio and X-ray maps of the remnant have been made with radio interferometers and X-ray telescopes and are shown in Colour Plate 19 (Reynoso *et al.*, 1997; Warren *et al.*, 2005).

These observations of Tycho’s remnant show both the shock in the ISM and the shocked ejecta. The forward shock is quite clear in both radio and X-ray images. The outer edge of the remnant is sharp and smooth. High-energy electrons moving in a region of relatively strong magnetic field produce the radio radiation and, as we shall discuss, most of the X-rays. The field is thought to be an interstellar field compressed and amplified by

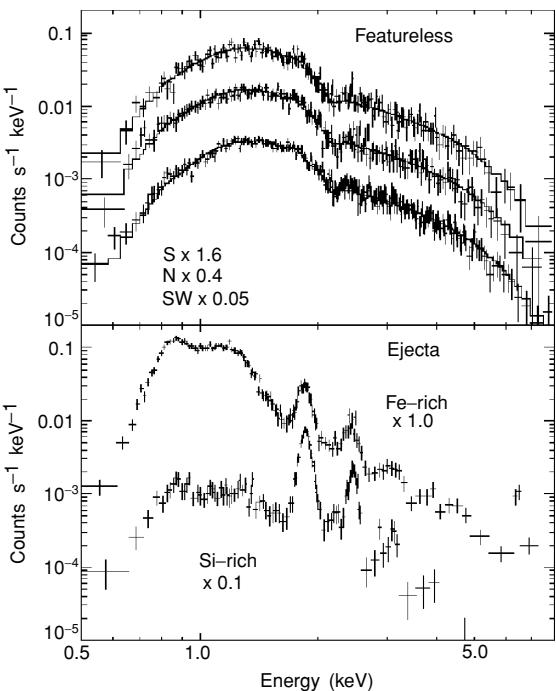


Fig 8.15 Chandra ACIS spectra from small sections of the remnant of Tycho's SN. (top) Spectra from three samples of the forward shock are all continuums (regions are indicated in Colour Plate 20; numbers are scaling factors for the display of each spectrum). (bottom) Spectra from Fe-rich and Si-rich regions. The Fe emission is mostly from L-shell transitions in the range 0.8–1.2 keV; Si emission is characterised by a K-shell line at 1.8 keV (Warren *et al.*, 2005).

the outer (or forward) shock (see Section 8.5 and Fig. 8.13).

8.6.1.2 X-ray structure

The ejecta show most clearly in X-rays. The brightest part of the remnant is a shell of material, broken into clumps, following the outer shock. Eighty per cent of the X-ray emission comes from this region. The composition of the material is revealed by the X-ray spectrum (Fig. 8.15) which is dominated by strong emission lines from Si and S. The inner part of the X-ray picture can be interpreted as a direct observation of material from the disrupted star. Deviations from spherical symmetry, other than the small-scale clumping, are not understood and may indicate an asymmetry in the explosion itself.

8.6.1.3 Modelling the X-ray spectrum reveals the progenitor

To understand the emission process and the arrangement of material, it is necessary to compare observations with a numerical model. This is not trivial. For one thing, because we are viewing a projection of a three-dimensional object, there is always an element of uncertainty in knowing the actual structure. In any case, the presence of lines in the X-ray spectrum identifies the emission process as thermal radiation from a hot gas (for SNR, a plasma of ions and electrons) and enables calculation of the temperature, composition and amount of material.

The deep Chandra and XMM-Newton observations of Tycho's SNR have been used to determine the distribution and motion of the shocked gas. The locations of the forward shock (FS), the contact discontinuity (CD) and the reverse shock (RS) were determined and are shown in Colour Plate 20 (Warren *et al.*, 2005). Emission from the region behind the forward shock is easily separated from that of the reverse-shocked ejecta. Emission lines from ions of Si, S, A and Ca in the ejecta are very strong, and because these elements were synthesised in the SN, characteristics of these lines can be used to determine the nature of the explosion itself (Badenes *et al.*, 2006). The result, after applying several models, was that the best fit is from a thermonuclear process which started as a subsonic burning (a deflagration) and then became supersonic (a detonation). This is called a ‘delayed detonation’. The analysis is sensitive to the explosive burning (which is finished in seconds) because the number of intermediate-mass elements and the amount of iron synthesised depend on the density of material when burned. Total amount of initial material was assumed to be that of a $1.4 M_{\odot}$ C-O white dwarf. The density of the surrounding material was assumed to be $\approx 1 \text{ atom cm}^{-3}$. The explosion model which fit the observations best synthesised $0.13 M_{\odot}$ of Si and $0.80 M_{\odot}$ of Fe, and the initial kinetic energy of the debris was $1.2 \times 10^{51} \text{ ergs}$ – an excellent fit to current ideas about the type Ia process.

The shock wave, clearly seen in both X-ray and radio images (Colour Plate 19), has properties that do not fit an adiabatic expansion. The average observed radial positions of RS:CD:FS are

Box 8.5 | Cosmic ray acceleration

The flux of *cosmic rays* at the top of Earth's atmosphere is $\approx 10^4$ particles $\text{m}^{-2} \text{ s}^{-1}$. The energy spectrum is a steep power law extending from $\approx 1 \text{ GeV}$ to $\approx 10^{20} \text{ eV}$, where the flux is ≈ 0.01 particles $\text{km}^{-2} \text{ yr}^{-1}$. The elemental composition is 'cosmic': 92 per cent protons, 7 per cent alpha particles and 1 per cent heavier ions. The electron flux is only a few per cent of the positive ion flux.

Where do these high-energy particles come from, and how are they accelerated? These are long-standing questions. A possible mechanism is *Fermi acceleration*, whereby high-energy particles are scattered from irregularly moving regions in the interstellar magnetic field. When the motion of a scattering centre is towards the particle (a head-on collision), the particle gains energy, and there is evidence now that this occurs efficiently in SNR shocks. A particle outside the shock passes through the shock, is scattered from high-velocity material behind the shock and gains energy. It then diffuses outside the shock and is scattered in again to repeat the process. Most of the collisions with scattering regions are head-on, as

opposed to 'following', and the process is called *first-order Fermi acceleration*. (In interstellar space, energy gain from head-on collisions is largely cancelled by energy loss from following collisions. However, it takes a particle longer to chase a scattering centre going in the same direction than to run into one head-on, so there are slightly more head-on collisions. This is *second-order Fermi acceleration*.)

The power-law X-ray spectra of forward shocks are the first indication that cosmic ray acceleration is common in young SNR, e.g. Figs. 8.15 and 8.16. The compressed interstellar field behind the shock should be $\approx 2 \times 10^{-5} \text{ G}$. Electrons of energy $\approx 10 \text{ GeV}$ in this field produce the non-thermal radio emission characteristic of almost all SNR. Because the frequency of synchrotron emission goes as $E^2 B$, 10^5 GeV electrons will radiate X-rays. So here is a source of 100-TeV electrons. There are also indications in young SNR that a large fraction of the energy behind the shocks is disappearing, e.g. Fig. 8.17. If this energy deficit goes into acceleration of both electrons and positive ions, this could account for the observed cosmic rays.

0.70:0.93:1.00 (Colour Plate 20) (Warren *et al.*, 2005). The CD is closer to the FS than predicted by most adiabatic models, e.g. 0.75:0.85:1.00 in Fig. 8.13. Furthermore, the CD is quite uneven and at some azimuths extends out to the FS. These deviations from adiabatic behavior are because of a loss of energy from behind the shock. Some kinetic energy of shocked material goes into particle acceleration. The particles acquire velocity $\approx c$ and soon leave the region, thereby removing energy from behind the shock. This results in less pressure behind the shock, which leads to a higher compression than the usual factor of 4 and to less space between the FS and CD. The greater compression also produces a peaked profile of the FS. These effects are clear in Figs. 8.15 and 8.17.

The bane of most SNR research is that the distance to remnants is often uncertain by a factor of ≈ 2 . In this case, the mass of material (which, if derived from the X-ray luminosity,

depends on the 5/2 power of distance) is uncertain by $\approx \times 6$.

The model used to fit the Tycho SNR observations assumes a progenitor of $1.4 M_{\odot}$ and a swept-up mass of $3.3 M_{\odot}$. The distance was treated as a variable and was derived to be 2.6 kpc, in good agreement with previous estimates. The remnant is no longer in the free expansion phase 1 and is starting the transition to phase 2.

As a final thought, consider the physical nature of this object. Although one of the smaller remnants, its 6-pc diameter is large compared to the distance scale we are used to. If the Sun were at the centre, our nearest celestial neighbours, the bright stars Alpha Centauri, Sirius and Procyon, and 12 faint red stars would all lie within the outer boundary of the remnant. The X-ray brightest clumps of material have a density of only $\approx 3 \text{ atoms cm}^{-3}$. Even though $4-5 M_{\odot}$ of material are within the outer boundary,

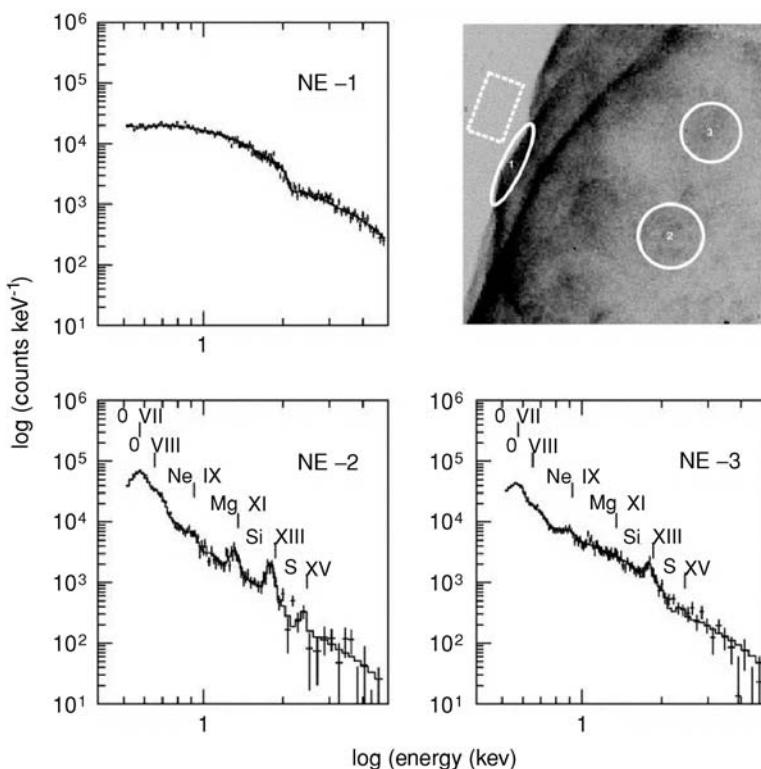


Fig 8.16 Chandra ACIS spectra from three regions of the remnant of SN 1006. The figure indicates where the three spectra were extracted and includes a box indicating background. Emission from the forward shock is a continuum. Emission from interior regions is softer and has lines characteristic of various ions, as labelled. Compare the image with the NE edge of Colour Plate 21, which shows the entire remnant (from Long *et al.*, 2003).

the remnant is completely transparent to radiation at all wavelengths. It is expanding into an almost pure hydrogen gas with density only $0.1\text{--}1 \text{ atoms cm}^{-3}$, a particle density $\approx 10^{-20}$ that of the air we breathe. The remnant is a very tenuous thing, truly a ghostly inhabitant of interstellar space.

8.6.2 The remnant of SN 1006 (type Ia)

8.6.2.1 Location

The historical location is only accurate to about 3° . It is, however, located 15° above the galactic plane, and there is only one radio source nearby with SNR characteristics: PKS 1459-41. X-rays were detected in 1971 (Palmieri *et al.*, 1972), and X-ray images and spectra were obtained by Einstein in 1980 (Pye *et al.*, 1981). The remnant is $30'$ in diameter, exactly the size of the full Moon, and the radio and X-ray images are almost identical. The Chandra image is shown in Colour Plate 21. The NE and SW limbs are bright. The NW limb is faint, and there is symmetry about an axis running NW–SE.

8.6.2.2 X-ray spectra

This remnant also shows clear evidence of cosmic-ray acceleration. The first spectra with the Einstein SSS instrument did not show prominent emission lines as did other young SNR. The possibility of synchrotron emission from high-energy electrons was mentioned (Becker *et al.*, 1980), but a thermal continuum also fit the data, and the community was not then ready to embrace the idea of cosmic-ray acceleration. In 1995, SNR 1006 was observed by ASCA, and it was shown that emission from the NE limb was a pure power-law continuum extending to 6 keV and the thermal emission was only from the interior. This ASCA result was accepted as proof of cosmic-ray acceleration in the forward shock of a SNR (Koyama *et al.*, 1995). The Chandra image in Colour Plate 21 shows this beautifully. Spectra from the shock and interior (Fig. 8.16) show that interior emission is soft, the strongest line is from ionised oxygen and the forward shock is a power-law continuum. Furthermore, the forward shock is strongly peaked, more so than the forward shock of Tycho's remnant (Long *et al.*, 2003). A large fraction of the energy behind

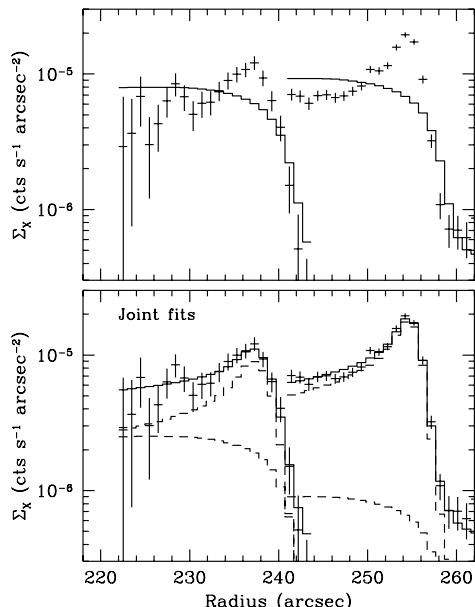


Fig 8.17 Radial profile of 4–6 keV X-ray surface brightness at the outer shock front of Tycho’s SNR. This shows the outer 15 per cent of two sectors of the remnant. (top) The solid curves show the profile expected from emission of hot gas behind an adiabatic shock. Because energy goes into acceleration of cosmic rays, the profile is peaked rather than flat. (bottom) The data fit with two components: a thin shell of emission and a lesser thermal component (from Warren *et al.*, 2005).

the shock is going into acceleration of particles. The spectacular SN has produced a remnant which is now a spectacular site of cosmic-ray acceleration.

8.6.3 Cassiopeia A (type II)

8.6.3.1 Location

Cas A is commonly regarded as the remnant of a type II supernova. The remnant is a bright ragged ring, as seen in radio and X-ray images, whose correspondence is good but not exact. Optical emission is faint and largely originates in small knots or clumps of material. The motion of optical fragments (Kamper & van den Bergh, 1976; Fesen *et al.*, 2006), if run backwards in time, implies an origin at about 1670, a time when there was a keen interest in astronomy in Europe. It is perplexing that this event, which occurred in a north circumpolar location, was not noted by anybody. Because type II supernovae are dimmer at maximum than

type Ia, a type II event has always been considered more likely. There are other unique features. The radio luminosity is very high; Cas A is the brightest radio source in the sky and was discovered by Grote Reber in 1943 (Reber, 1944). The interior of the remnant is full of bright X-ray and optical knots; some are highly enriched in O or Fe. The remnant seems much more clumpy than other young remnants.

8.6.3.2 X-ray properties

Colour Plate 22 shows a very deep (1 million second) Chandra observation of Cas A, from which the global distribution of metals in the remnant can be mapped. The signal-to-noise ratio is so good that abundances and ion states of the material in each knot can be determined. With this information the progress of the reverse shock through the interior can be followed and asymmetries in the explosion measured. The forward shock is marked by the 4–6 keV continuum. This is quite patchy but circular in projection and well outside the debris illustrated here by Si and Fe emission. The debris seems contained in the interior but for jets extending beyond the FS in the NE and SW. The NE jet is considerably brighter than the SW jet. Figure 8.18 shows spectra from the two jets. These are typical of the ejecta in general, but there is a great deal of variation in individual knots within the remnant. The figure also shows the spectrum of a region dominated by Fe. Indications are that Cas A contains $2 M_\odot$ of ejecta and that the energy released in the SN explosion was 2×10^{51} ergs.

A faint compact X-ray source at the centre of the remnant verifies that this was a core-collapse SN. This pointlike source, accounting for only a small fraction (3×10^{-4}) of the 0.5–10 keV emission, was not known until the Chandra first-light observation, a splendid demonstration of the value of high spatial resolution, the finest ever in X-ray astronomy.

8.6.4 SNR 0102-72.3 in the Small Magellanic Cloud (type II)

The brightest soft X-ray source in the Small Magellanic Cloud is a SNR discovered by Einstein in 1980 (Seward & Mitchell, 1981). The Chandra ACIS spectrum is shown in Fig. 8.19, and

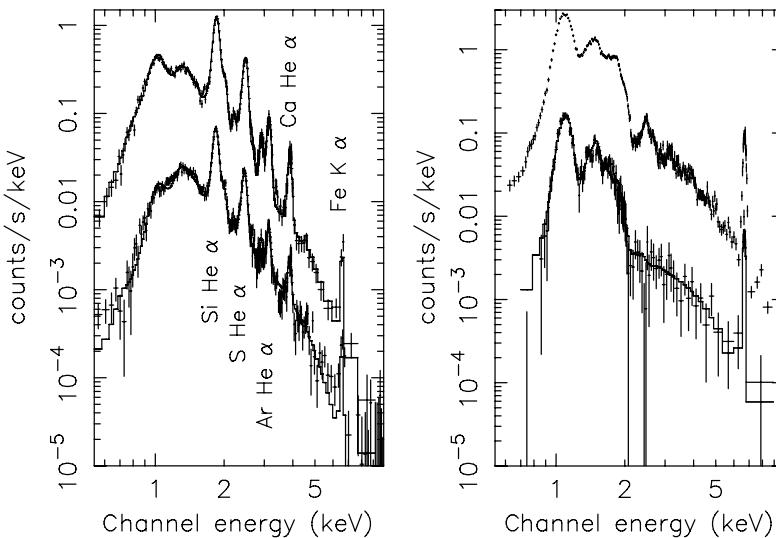


Fig 8.18 Chandra X-ray spectra of small regions within Cas A. (left) Bright filaments in the jets; upper curve is from NE jet and lower curve is from SW jet. There are bright emission lines from several elements. Relative amounts are different in the two regions. (right) Emission from a cloud in the SE shell which is apparently close to pure Fe. Upper curve is the spectrum obtained in 2004 and lower curve is the spectrum from 2002 observations. There is strong Fe L- and K-shell emission (from Hwang *et al.*, 2004).

is quite soft, with dominant oxygen and neon emission lines. These spectra show that, also in this remnant, the spectrum of the FS is a continuum, and emission lines come only from interior reverse-shocked ejecta. The remnant is only 40 arcsec across, which makes it an excellent target for dispersive spectroscopy. Two observations totalling 135 ks with the Chandra HETG grating produced the dispersed image shown in Fig. 8.20. Each emission line produces an image of the remnant at a dispersion determined by

the line energy. Each dispersed image shows the distribution of the emitting element within the remnant. The forward shock shows only in the undispersed, or zero-order, image, as is clearly visible in Fig. 8.21.

The relative numbers of H-like and He-like ions of a given element are determined by the temperature and ionisation timescale (Chapter 2 and Box 8.4). These properties and total amounts of the various elements can be calculated. Indications are that $\approx 6 M_{\odot}$ of O and $\approx 2 M_{\odot}$ of Ne are

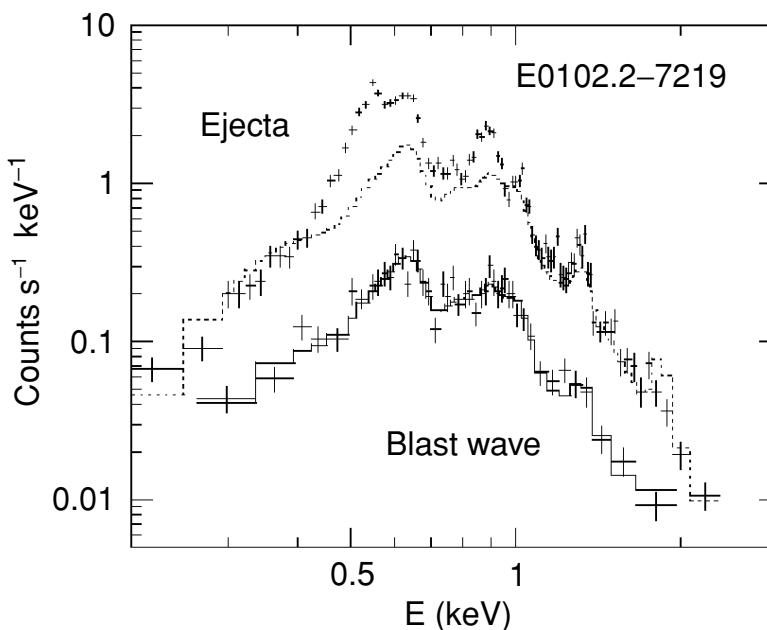


Fig 8.19 Chandra ACIS spectra of two portions of SNR 0102-72.3. Exposure time was 9 ks. Upper data were taken from inner part of the shell where the reverse shock has heated the ejecta. The lower spectrum is from the outer part of the shell where there is only shock-heated circumstellar material. Dashed curves show a blast wave model fitted to the two spectra. The ejecta have emission lines from O and Ne that are not seen in the blast wave spectrum (from Hughes *et al.*, 2000).

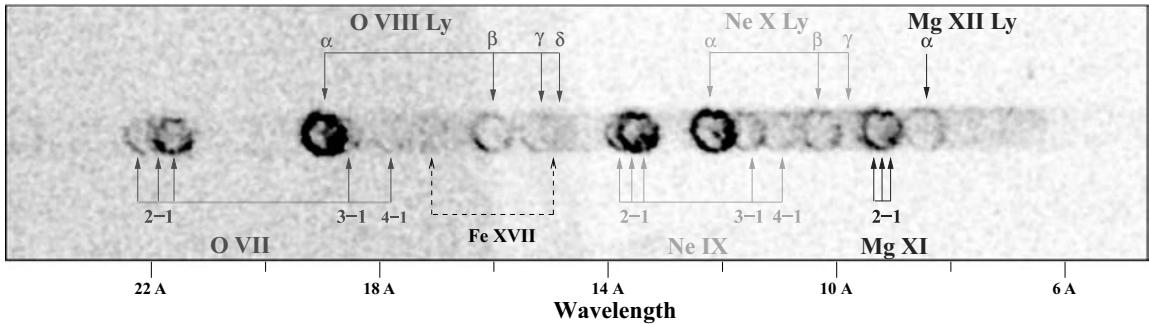


Fig 8.20 The dispersed spectrum of SNR 0102-72.3 obtained by the Chandra HETG/ACIS instrument in two exposures totaling 138 ks. The dispersed -1 order is shown. The 0 order is not. Transitions of several ion species are indicated. Ne and O species are strong. Fe emission is weak (from Flanagan et al., 2004).

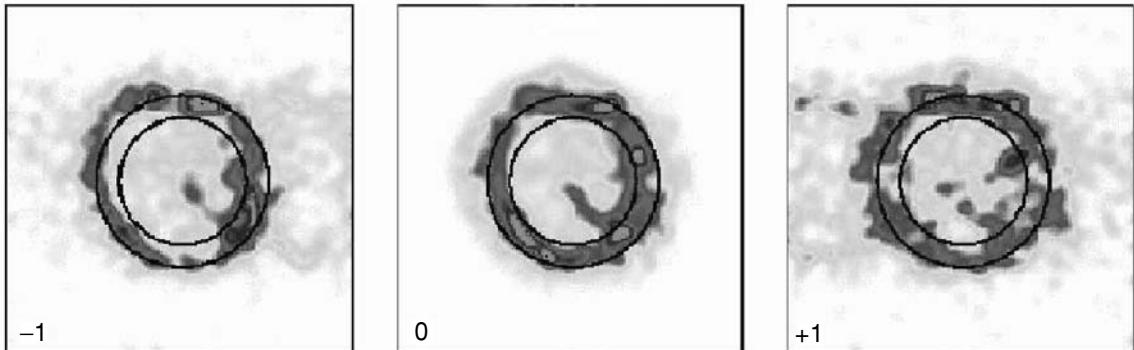


Fig 8.21 Dispersed images from Chandra FPCS/MEG. The Ne X Ly α line from SNR 0102-72.3 in -1 , 0 and $+1$ orders. These show distortion caused by Doppler shifts. The left ring is slightly squeezed and the right ring slightly stretched in the direction of dispersion, showing that the velocity of material in the SE limb is directed somewhat towards the observer (from Flanagan et al., 2004).

present. Comparison with core-collapse SN models favours a $30 M_{\odot}$ progenitor.

The Chandra HETG is described in Chapter 3. For monochromatic radiation, there is, for each grating, an undispersed zeroth-order image and two dispersed first-order images; one to the left, the ' -1 order', and one to the right, the ' $+1$ order'. If the radiating source is at rest, the dispersion distance is the same for $+$ and $-$ orders. The SNR 0102-72.3 source approximates to a ring and is expanding. The east (left) edge is moving away from us, so there is a red shift, and the dispersed image of this edge is pushed away from the zeroth order. The west (right) edge is moving towards us, so there is a blue shift, and the dispersed image of this edge is pulled towards the zeroth order. Thus the -1 order image is slightly elongated along the dispersion direction and the $+1$ order is slightly

compressed. Figure 8.21 shows this small but definite effect in the NeX Ly α images. The velocity derived is $450 \pm 100 \text{ km s}^{-1}$. These data have been used to generate a geometric model which consists of a non-uniform spherical shell with emission concentrated towards the equatorial plane (Flanagan et al., 2004). The same techniques have been applied to other remnants, but they are generally larger and not as bright as SNR 0102-72.3, which makes analysis difficult.

8.7 Dominated by their surroundings: Older remnants

Interstellar gas is swept up as remnants age and become massive and cooler. Rather dense

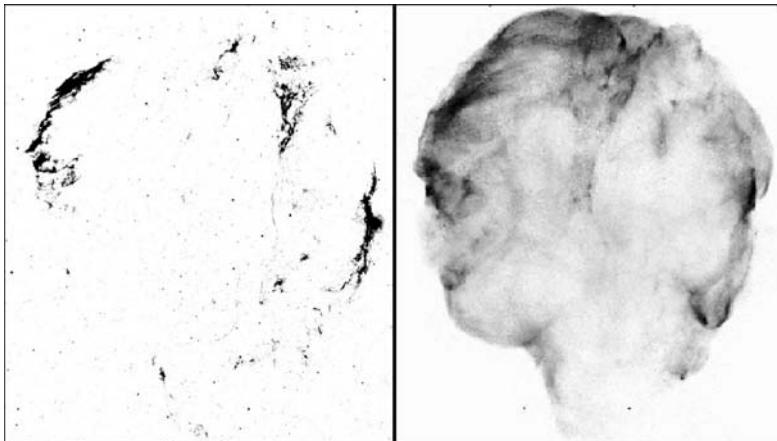


Fig 8.22 (left) Optical and (right) X-ray images of the Cygnus Loop. The remnant has a diameter of 2.6° with a prominent 'blowout' in the south. The X-ray image is from the ROSAT HRI and required a mosaic of 47 exposures (totalling $\approx 10^6$ s) to cover the remnant. Energy range is 0.1–2.4 keV (Levenson *et al.*, 1997, via HEASARC archive).

filaments can form which radiate efficiently in the optical and UV bands. Many older remnants are visually prominent and have been studied extensively using optical observations. Because space is short and this is a book on X-ray astronomy, the presentation of data at other wavelengths is limited. We note, however, that many of the basic properties of older remnants are derived from optical (or radio) data and emphasise this where appropriate.

8.7.1 The Cygnus Loop

The Veil Nebula, NGC 6992, is a well-known, prominent optical feature at the southern edge of the constellation Cygnus. Together with smaller structures (NGC 6979, 6990), all falling approximately on the circumference of a 2.5° diameter circle, it forms a structure known as the Cygnus Loop. A radio map verified that this was indeed a SNR with the optical filaments concentrated along the limb. This was identified as an X-ray emitter in 1968 (Grader *et al.*, 1970) and shows clearly as a bright, soft, extended source in early slit-collimator rocket data, as in Fig. 1.6.

Because the remnant has a large angular extent, it is time consuming to map with X-ray telescopes with only arcminute fields of view. Because the X-rays are soft, the ROSAT map in Fig. 8.22 is one of the best. X-ray and radio morphologies are similar, but intensity variation around the shell differs appreciably. This is a classic shell-like remnant. The prominent shell

is not an indication of the forward shock. It is a circumstellar structure caused by the variation of wind mass loss from the progenitor star. During the progenitor's evolution, immediately prior to the SN explosion, a massive slow wind changed to an energetic high-velocity wind. The fast wind swept material from the slow wind before it and pushed it into a relatively dense shell. This shell has now been heated by the forward shock and radiates soft X-rays.

The forward shock itself can be seen in the light of optical H α and lies just outside the outermost X-ray emission (Fig. 8.23). Molecular clouds at the east and west limbs (three o'clock and nine o'clock) have slowed the shock and caused indentations in the shell. The resultant reverse shock has heated these clouds, and emission is bright at the indentations. Figure 8.24 shows X-ray spectra from FS gas and RS cloud, which are dominated by oxygen. There is also a rich array of filaments seen in optical lines of hydrogen, plus oxygen and sulphur ions, from which temperature, velocity and structure can be derived.

8.7.2 Mixed-morphology remnants

Figure 9.13 shows the mixed-morphology remnant Kes 79. Radio contours are overlaid on a grey-scale X-ray map. As with other members of this class, the outer shell is prominent in the radio map but weak or absent in X-rays. Typically, there is strong diffuse X-ray emission from the central regions, and the X-ray spectrum is thermal. There is little variation in the temperature from place to place,

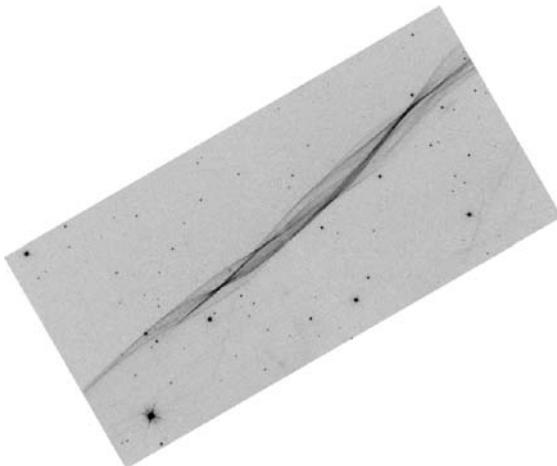


Fig 8.23 HST observation of H α emission from the forward shock just outside the Veil Nebula in the NE part of the Cygnus Loop. The shocked material forms a very thin rumpled sheet which is here viewed edge on. The field is 1.3' \times 2.7', north is up and the shock is travelling outward, to the NE (from Blair et al., 2005).

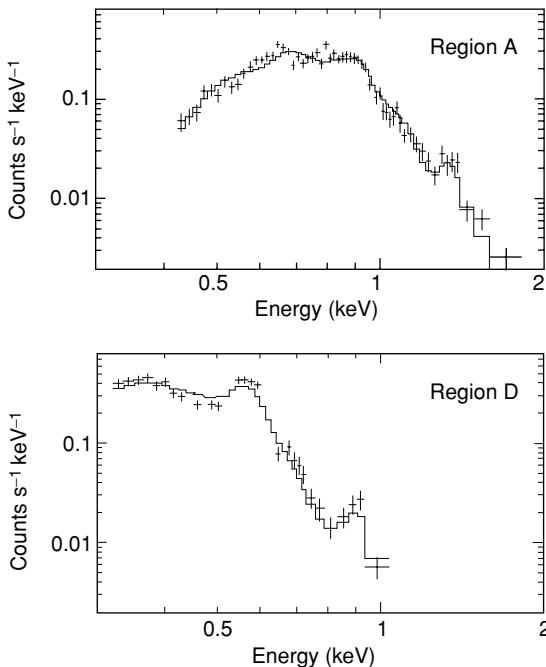


Fig 8.24 Chandra ACIS spectra of small regions along the west limb of the Cygnus Loop. (top) (region A) Emission from the reverse shock in a molecular cloud. (bottom) (region D) Extracted from just behind the forward shock and showing emission from O and Ne (from Levenson et al., 2002).

and it is thought that thermal conduction is high and, therefore, quickly smooths out any large temperature fluctuations. These remnants are middle-aged with ages of a few thousand years. Several have faint central neutron stars such as the one seen in Figure 9.13.

Box 8.6 Classification of supernova remnants

From an X-ray perspective, SNR are sometimes classified according to morphology. There are four categories:

shell-like. Both X-ray and radio images are strongly limb brightened. In older remnants, this may not extend over the entire limb. Examples are the remnant of Tycho's SN and the Cygnus Loop.

filled-centre (or *Crab-like* or *plerionic*). Both radio and X-ray images are brightest at the centre. Limb brightening is absent or very weak. Examples are the Crab Nebula and 3C 58.

composite. Radio and X-ray images show both limb brightening and bright centres. Usually the central emission is because of a PWN. Examples are Kes 75 and SNR 0540-69.3.

mixed morphology (or *thermal composite*). These remnants appear shell-like in the radio band but, in the X-ray band, have filled centres. X-rays from the interior are usually found to be thermal. If there is a central PSR/PWN complex, it is weak. Examples are W44, IC 443 and Kes 79.

Properties of some well-known remnants are listed in Table 8.2.

Two mechanisms to produce this morphology have been proposed. The first assumes that the interior contains many moderately dense interstellar clouds. These are heated by the shock, and material is evaporated to fill the space within the remnant. There are many indications of clouds at the rim, but those in the interior have been hard to spot (White & Long, 1991). The second explanation is that the interior gas is that which is expected from the Sedov model and the cooler gas at the shocked shell is not seen because the softer X-rays are absorbed in the ISM. Most mixed-morphology remnants are fairly distant and located in the galactic plane, so absorption is high.

Table 8.2 Some well-studied mostly thermal emission SNRs.

Name	Age (kyr)	Dist. (kpc)	Diam. (pc)	Angular diam. (arcmin)	Phase	Type	Einstein IPC rate (count s ⁻¹)	L_X (10^{35} ergs s ⁻¹)
Cas A	0.34	3	3.5	4	1–2	Shell-like	61	38
Kepler	0.40	5?	4.4?	3	1–2	Shell-like	7	10
Tycho	0.43	3	7.0	8	1–2	Shell-like	22	6
SN 1006	1.00	1	9	30	1–2	Shell-like	11	1
SNR 0102-72.3	1	60	13	0.73	1–2	Shell-like	0.8	210
W44	~7	3	≈26	≈30	2	Mixed morph.	3.3	20
Kes 79	~7	7	22	11	2	Mixed morph.	0.24	25
Vela XYZ	~10	0.3	29	330	2–3	Composite	500	2
Cyg Loop	~20	0.8	40	170	2–3	Shell-like	620	11

Neither theory produces good fits to the shape of these remnants. There are large clouds nearby in some cases, and the interaction with a large cloud explains the gross properties of the remnant. The material in older remnants is almost all interstellar, so it is not surprising that the morphology eventually reflects the structure of the ISM in the vicinity of the progenitor.

8.8 | Summary

There are now ≈ 270 known SNR in the Galaxy, ≈ 50 in the Magellanic Clouds and appreciable numbers are being discovered in M31, M33 and other nearby galaxies. They are all individuals; no two look the same. A few have been presented in this chapter to give a sense of what can be learned. Clearly remnants start out small and hot and expand and cool with time. There are clearly forward shocks, and much of the physics can be derived from their properties. The application of simple models is limited by non-uniformity of circumstellar material, instabilities in expanding debris and non-equilibrium of the radiating material. Nevertheless our ideas about progenitor stars, the supernova process and properties of the remnants are in excellent agreement. Recent observations have produced a mountain of detailed information. A full understanding will take a while.

References

- Badenes, C., et al. (2006) Constraints on the physics of type Ia supernovae from the X-ray spectrum of the Tycho supernova remnant, *ApJ* **645**, 1373.
- Becker, R. H., et al. (1980) Is the remnant of SN 1006 Crablike?, *ApJ* **240**, L33.
- Blair, W., R. Sankrit & J. Raymond (2005) Hubble Space Telescope imaging of the primary shock front in the Cygnus Loop supernova remnant, *ApJ* **129**, 2268.
- Blondin, J. M. & D. C. Ellison (2001) Rayleigh-Taylor instabilities in young supernova remnants undergoing efficient particle acceleration, *ApJ* **560**, 244.
- Cassam-Chenai et al. (2008) Morphological evidence for azimuthal variations of the cosmic-ray ion acceleration at the blast wave of SN1006, *ApJ* **680**, 1180.
- Chevalier, R. A. (1982) Self-similar solutions for the interaction of stellar ejecta with an external medium, *ApJ* **258**, 790.
- Colgate, S. A. (1974) Early gamma rays from supernovae, *ApJ* **187**, 333.
- Cox, D. P. (1972) Cooling and evolution of a supernova remnant, *ApJ* **178**, 159.
- Dotani, T., et al. (1987) Discovery of an unusual hard X-ray source in the region of supernova 1987A, *Nature* **330**, 230.
- Fesen, R. A., et al. (2006) The expansion asymmetry and age of the Cassiopeia A supernova remnant, *ApJ* **645**, 283.
- Flanagan, K. A., et al. (2004) Chandra high-resolution X-ray spectrum of supernova remnant 1E 0102.2-7219, *ApJ* **605**, 230.

- Friedman, H., E. T. Byram & T. A. Chubb (1967) Distribution and variability of cosmic X-ray sources, *Science* **156**, 374.
- Gorenstein, P., E. Kellogg & H. Gursky, (1970) X-ray characteristics of three supernova remnants, *ApJ* **160**, 199.
- Grader, R. J., R. W. Hill & J. P. Stoering (1970) Soft X-rays from the Cygnus Loop, *ApJ* **161**, L45.
- Hanbury-Brown, R. & C. Hazard (1952) Radio-frequency radiation from Tycho Brahe's supernova (A.D. 1572), *Nature* **170**, 364.
- Hughes, J., C. Rakowski & A. Decourchelle (2000) Electron heating and cosmic rays at a supernova shock from Chandra X-ray observations of 1E 0102.2-7219, *ApJ* **543**, L61.
- Hwang, U., et al. (2004) A million second Chandra view of Cassiopeia A, *ApJ* **615**, L117.
- Immler, S., B. Aschenbach & Q. D. Wang (2001) X-ray detection of presupernova evolution for the SN 1993J progenitor, *ApJ* **561**, L107.
- Inoue, H., et al. (1991) X-ray observation of SN 1987A from GINGA, *PASJ* **43**, 213.
- Kamper, K. & S. van den Bergh (1976) Cassiopeia A – an unseen supernova, *Sky Telescope* **51**, 236.
- Koyama, K., et al. (1995) Evidence for shock acceleration of high-energy electrons in the supernova remnant SN1006, *Nature* **378**, 255.
- Levenson, N. A., et al. (1997) The ROSAT HRI Survey of the Cygnus Loop, *ApJ* **484**, 304.
- Levenson, N. A., J. R. Graham & J. L. Walters (2002) Shell shock and cloud shock: Results from spatially resolved X-ray spectroscopy with Chandra in the Cygnus Loop, *ApJ* **576**, 798.
- Long, K., et al. (2003) Chandra CCD imagery of the northeast and northwest limbs of SN 1006, *ApJ* **586**, 1162.
- NASA/CXC/GSFC (2004) <http://chandra.harvard.edu/photo/2004/casa>.
- NASA/CXC/Rutgers (2005) <http://chandra.harvard.edu/photo/2005/sn1006>.
- Palmieri, T. M., et al. (1972) Observations of soft X-rays: Two supernova remnants in the constellation Lupus and the diffuse background, *ApJ* **177**, 387.
- Park, S., et al. (2006) Evolutionary status of SNR 1987A at the age of eighteen, *ApJ* **646**, 1001.
- Perlmutter, S. & B. Schmidt (2003) Measuring cosmology with supernovae, in *Supernovae and Gamma-ray Bursters*, edited by K. Weiler, 195, Berlin: Springer.
- Pye, J. P., et al. (1981) An X-ray map of SN 1006 from the Einstein Observatory, *MNRAS* **194**, 569.
- Reber, G. (1944) Cosmic static, *ApJ* **100**, 279.
- Reynoso, E. M., et al. (1997) A VLA study of the expansion of Tycho's supernova remnant, *ApJ* **491**, 816.
- Rho, J. H., et al. (1994) An X-ray and optical study of the supernova remnant W44, *ApJ* **430**, 757.
- Seward, F. D. & M. Mitchell (1981) X-ray survey of the Small Magellanic Cloud, *ApJ* **243**, 736.
- Seward, F., et al. (2003) A compact central object in the supernova remnant Kes 79, *ApJ* **584**, 414.
- Soderberg, A. M., et al. (2008) An extremely luminous X-ray outburst at the birth of a supernova, *Nature* **453**, 469.
- Sunyaev, R., et al. (1988) Hard X-ray brightening of SN 1987A – The Mir-Kvant data, *SvAL* **14**, 288.
- van den Bergh, S. (1971) The expansion of the optical remnant of B Cassiopeiae = 3C 10, *ApJ* **168**, 37.
- Warren, J. S., et al. (2005) Cosmic-ray acceleration at the forward shock in Tycho's supernova remnant: Evidence from Chandra X-ray observations, *ApJ* **634**, 376.
- White, R. & K. Long (1991) Supernova remnant evolution in an interstellar medium with evaporating clouds, *ApJ* **373**, 543.
- Zhekov, S. A., et al. (2005) Chandra observations of shock kinematics in supernova remnant 1987A, *ApJ* **628**, L127.

Catalogues, reviews, popular

- Brown, G. & H. A. Bethe (1985) How a supernova explodes, *Sci Am* **252**, 60.
- Clark, D. H. & F. R. Stephenson (1977) *The Historical Supernovae*, Oxford: Pergamon Press.
- Green, D. A., (2006) A catalogue of galactic supernova remnants, <http://www.mrao.cam.ac.uk/surveys/snrs/>.
- Seward, F. D., et al. (2007) Chandra supernova remnant catalog, <http://hea-www.cfa.harvard.edu/ChandraSNR/>.
- Soderberg, A. (2008) A newborn supernova, *Sky Telescope* **116**, 26.
- Weiler, K. & R. Sramek (1988) Supernovae and supernova remnants, *Annu. Rev Astron Appl* **26**, 295.

Chapter 9

Neutron stars, pulsars, pulsar wind nebulae, and more supernova remnants

9.1 Discovery and nature of neutron stars

Until the discovery of radio pulsars by Jocelyn Bell and Antony Hewish in 1967 (Hewish *et al.*, 1968), neutron stars had existed only in the minds of theoretical physicists. First proposed as an end state of stellar evolution by Robert Oppenheimer and George Volkoff (1939), they are now accepted as the only explanation for radio pulsars. The discovery was serendipitous. No one had conjectured or even dreamed that this sort of signal might be generated.

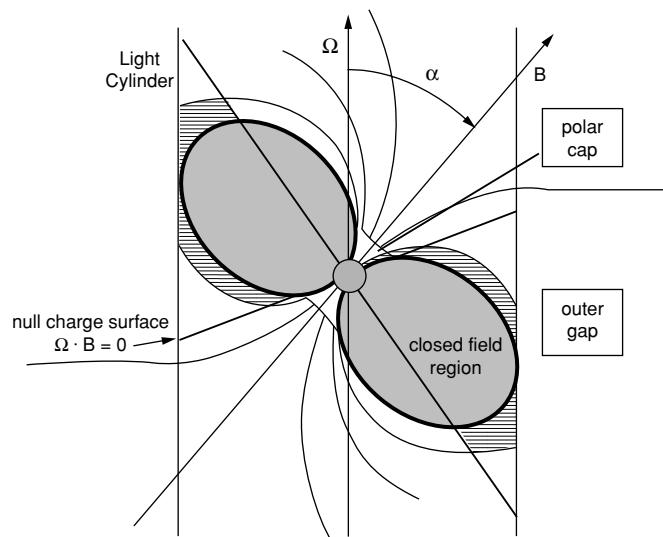
Bell, then a student, had just spent months wiring antennas for a new radio telescope. In the course of testing, she noticed ‘a bit of scruff’ in the recorded signal. This ‘scruff’ was found to repeat, not every 24 hours (the *solar day*), but every 23 hours and 56 minutes (the *sidereal day*), showing that the source was anchored in the sky, not to the rotating Earth. Furthermore, when present, the signal was periodic with a remarkably stable period of 1.337 s. Thus CP 1919, the first of the radio pulsars, was discovered. The stable period could only be explained by rotation, and only a small object with strong gravity could rotate this fast without breaking up. A neutron star was the only reasonable explanation (Gold, 1968).

The neutron star has a strong magnetic field which, to a first approximation, is a dipole (like a bar magnet). This field probably originated in the progenitor star and was compressed and concentrated by the gravitational collapse at the time

of the supernova which produced the neutron star. Like the Earth, the magnetic dipole is not aligned with the axis of rotation. At the surface, the magnetic poles are in different places than the poles of rotation (see Fig. 9.1). As the star rotates, a pulse of radiation is observed every time one of the magnetic poles passes through our field of view. Depending on orientation, an observer sees one or two pulses every rotation. Long-term radio monitoring soon showed that CP 1919 was slowing down; the period of the pulsar was very slowly, but very steadily increasing at a rate of $0.042 \mu\text{s}$ per year. The loss of rotational energy that this slowing implies was more than enough to power the observed radiation. Almost all the more than 1700 known radio pulsars exhibit this behavior.

The neutron star, from our perspective, is a bizarre state of matter. A mass about equal to that of the Sun is confined to a region 15–30 km in diameter. The material is almost all neutrons. The density is that of the atomic nucleus, $\sim 10^{14} \text{ g cm}^{-3}$, perhaps $10^{15} \text{ g cm}^{-3}$ at the centre. A 10-km cube of rock (Mt Everest) compressed to a 10 cm cube (a grapefruit) would have this density. The gravitational field close to the surface is enormous, and the tidal force will pulverise to dust any solid object which approaches closer than a few thousand kilometers. We know these stars exist and we know the masses of a few which are in binary systems (see Chapter 11). There is no straightforward measure of the radii or the internal structure. We would dearly like to know more. The only process which can create such an object, which can squeeze matter until all the electrons

Fig 9.1 Rotating magnetised neutron star and its magnetosphere. Magnetic field is that of a dipole inclined 40° to the vertical rotation axis. Proposed regions for generation of pulsed radiation are indicated (adapted from Kaspi et al., 2006).



and protons merge into neutrons, is the gravitational collapse of the core of a massive star, the event which triggers a type II (or type Ib/Ic) supernova.

9.2 | The Crab Nebula

The most famous, the most studied type II supernova remnant, the one dear to the heart of most astrophysicists, is the Crab Nebula. It contains the first-discovered and undisputed example of a neutron star formed in a supernova explosion. The evidence for this is overwhelming: the supernova was observed in 1054 and its location is the same as that of the remnant; the observed expansion of the remnant, run backwards in time, converges at about the date of the supernova; the pulsar is located close to the centre of the remnant; the characteristic age of the pulsar is close to the age of the remnant; and the rate of rotational energy loss of the pulsar is high enough to support the high luminosity of the Crab Nebula at all wavelengths. Here then is a remnant known to have originated in gravitational collapse which created a neutron star! Colour Plate 23 shows the Crab Nebula at X-ray and optical wavelengths.

The discovery of the Crab pulsar by David Staelin and Edward Reifenstein (1968) solved a long-standing problem of astrophysics. The

pulsar was the powerhouse supplying the prodigious energy radiated by the nebula. The total energy output of the Crab Nebula, 900 years after the supernova explosion, from radio to X-rays is an incredible 1.8×10^{38} ergs s $^{-1}$, about 100 000 times the energy output of the Sun. With a period of 33 m, this pulsar was, for many years, the fastest known pulsar. There are two bursts of radiation each cycle (Fig. 9.2), and the pulse shape is about the same throughout the electromagnetic spectrum. Because of the central injection of energy by the pulsar, the Crab is not shell-like but 'filled in' in appearance. Hence the term *plerions* (from the Greek word *pleres*, meaning 'full') is sometimes used to describe Crab-type remnants.

The intense low-frequency electromagnetic radiation of the pulsar (a giant rotating bar magnet) produces a torque which slows the rotation. A large fraction of the energy in this low-frequency radiation is transferred quickly into high-energy electrons and magnetic field. These electrons form a 'relativistic wind' flowing outwards from the pulsar. The exact mechanism by which electrons are accelerated to high energies is, alas, not understood. In any case, the pulsar is surrounded by a cloud of relativistic particles, and when the electrons cross magnetic field lines, they radiate synchrotron radiation, forming a bright, diffuse nebula. This is called the *synchrotron nebula* or *pulsar wind nebula* (PWN). The Crab PWN radiates a

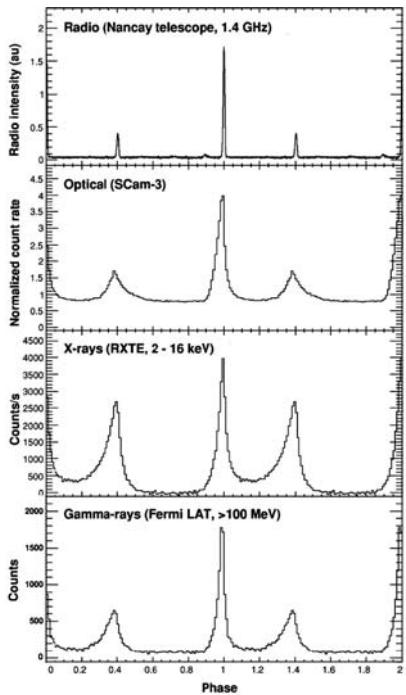


Fig 9.2 Pulsed radiation from the Crab pulsar as observed in four wavebands from radio to high-energy γ rays. The curves are labelled with instrument, observatory and energy range. It is remarkable that the double-pulse structure occurs at the same phase throughout the electromagnetic spectrum (adapted from Abdo *et al.*, 2010).

continuum throughout the electromagnetic spectrum from radio waves to γ rays, as shown in Fig. 9.3. It dominates the appearance of the Crab at X-ray wavelengths, and radiation from this nebula is quite separate from the pulsed radiation which originates close to the spinning neutron star.

X-rays from the Crab Nebula were first detected with a rocket-borne detector on 12 October 1962 (Gursky *et al.*, 1963). The Crab was located well enough to identify in data from a flight on 29 April 1963 (Bowyer *et al.*, 1964a) and confirm during a 7 July 1964 lunar occultation (Bowyer *et al.*, 1964b). After a very careful calculation, a rocket was launched by Stuart Bowyer and colleagues and the detector pointed at the nebula. The relative motion of the Earth, the Moon and the rocket then caused the nebula to pass behind the limb of the Moon; all during the 5-min flight of the rocket! The emitting region was not only precisely located but was established as being extended.

The next series of lunar occultations occurred in 1974. During these the Crab Nebula was observed by several rockets and balloon-borne payloads and the shape and size of the X-ray nebula were determined. The pulsar itself appeared clearly as a pulsing point source, but to our surprise, the synchrotron nebula was not symmetrical around the pulsar. The brightest X-ray emission was centred 10 arcsec to the northwest.

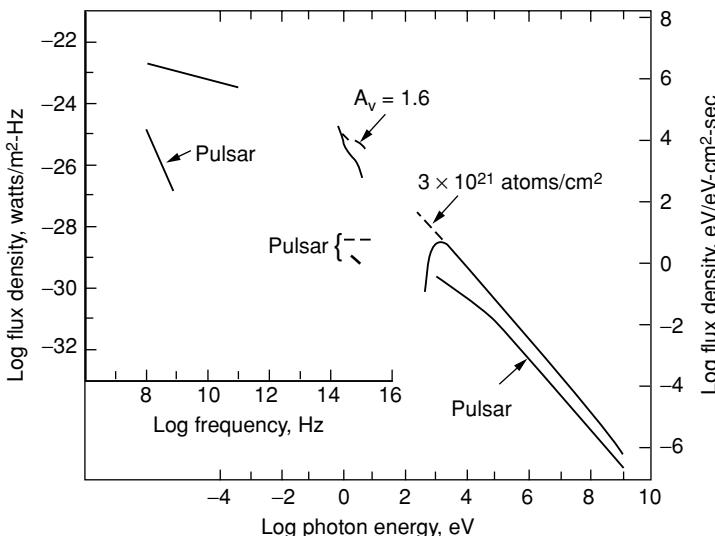


Fig 9.3 Electromagnetic spectra of the Crab Nebula and of the Crab pulsar. Scales at left show radio units and at right X-ray units. Upper curve is the diffuse PWN. Lower curve is the pulsar. Dashed lines show corrections for X-ray absorption and optical extinction in interstellar material. Spectra are power-law continua and are well measured. There is no structure. Note the change in the PWN spectral index in the infrared and the strength of the pulsar at the lowest frequencies measured (from Seward, 1978).

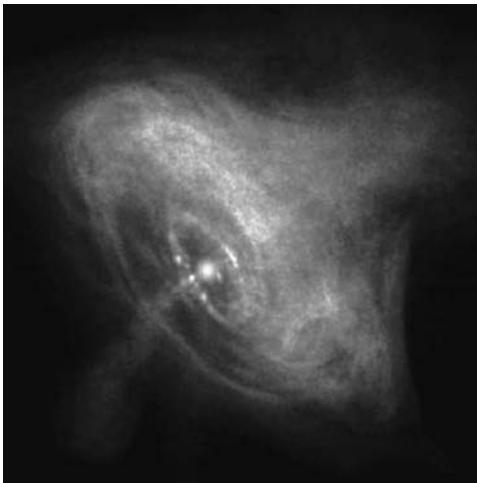


Fig 9.4 The innermost bright region of the Crab PWN. This Chandra X-ray snapshot (7-hour exposure) shows the complexity of the PWN structure. The central pulsar is surrounded by a torus of bright, outward-moving rings. The count rate is very high from the pulsar itself, and the detector saturates. This image is ≈ 2 arcmin square (from Hester *et al.*, 2002).

Einstein and ROSAT observations in 1979 and 1990 produced clear X-ray images, showing that the dominant feature was a torus surrounding the pulsar, slightly inclined and brightest on the side facing the observer (Harnden & Seward, 1984; Brinkmann *et al.*, 1985). A Chandra observation in 1999 produced the image shown in Fig. 9.4 (Weiskopf *et al.*, 2000; Hester *et al.*, 2002).

The pulsar is surrounded by bright rings, and there is fine structure within the torus. There is a distinct jet extending southeast from the pulsar. Because it is nearby and energetic, the Crab is our best example of the processes that occur in SNR/PSR/PWN systems. We explain these somewhat in the next sections but only touch the surface of a vast body of observation and theory.

9.3 | Rotation and spindown

A rotating magnetic dipole with period P must radiate electromagnetic energy at the rotation frequency $\Omega = 2\pi/P$. There is a torque associated

with this radiation which acts to slow the rotation. Thus rotational energy is changed to radiation, and as time advances, the star spins more slowly. It is assumed that the torque is proportional to Ω^n , where n is the *braking index*. If the field is a dipole field and the axis of the dipole is perpendicular to the rotation axis, this torque is at a maximum. Several quantities can be simply calculated from the observed P (measured in seconds) and the spindown rate (or period derivative) \dot{P} . The rotational moment of inertia I is calculated from one of several proposed equations of state of nuclear material and, for a $1.4 M_\odot$ neutron star, usually has a value close to $10^{45} \text{ g cm}^2 \text{ s}^{-2}$. R is the neutron star radius and is about 10 km. The rotational energy and the rate of loss of this energy are as follows:

$$E = 2\pi^2 I / P^2 \text{ ergs} \quad \dot{E} = -4\pi^2 I \dot{P} / P^3 \text{ ergs s}^{-1}.$$

Since the pulsar is losing energy, \dot{E} is negative. Assuming the pulsar is born with initial period P_0 and the magnetic field does not change with time, the age τ of the pulsar is

$$\tau = \frac{P}{(n-1)\dot{P}} \left[1 - \left(\frac{P_0}{P} \right)^{n-1} \right].$$

If the pulsar has slowed appreciably and the period is now much larger than the period at birth, the age of the pulsar should be close to the ‘characteristic age’,

$$\tau \approx A = P / 2\dot{P} \text{ s}.$$

B is the magnetic field at the pole where it is strongest and, in the case of pulsars, is very strong indeed:

$$B = (3c^3 I P \dot{P} / 8\pi^2 R^6)^{1/2} \text{ G}.$$

The field of the Crab Pulsar is calculated to be 5×10^{12} G at the poles—stronger than any field in our experience. The natural magnetic field at the Earth’s surface is strongest at the south magnetic pole, where it is 0.7 G. The strongest steady field we have been able to make in a laboratory is 3×10^5 G. Materials available on Earth do not have the strength to contain stronger fields. Consider, for example, the energy density of the field ($B^2/8\pi$) at the pole of the Crab pulsar, which is 10^{24} ergs cm^{-3} , equivalent to 25 Mtons of TNT in each cm^3 !

Table 9.1 Properties of selected pulsars.

Pulsar	P	\dot{P} ($s s^{-1}$)	\dot{E} (ergs s^{-1})	A (kyr)	B (G)	n	Distance (kpc)	Comment
CP 1919	1337 ms	1.34×10^{-15}	2.2×10^{31}	1.6×10^4	1.3×10^{12}	0.7	Radio PSR	
PSR 1257+12	6.2 ms	1.21×10^{-19}	2.0×10^{34}	8×10^5	9×10^8	0.8	ms PSR	
Crab pulsar	33 ms	4.22×10^{-13}	4.6×10^{38}	1.24	3.8×10^{12}	2.51	2.0	Young PSR
Vela pulsar	89 ms	1.24×10^{-13}	7.1×10^{36}	11.3	3.4×10^{12}	1.4	0.29	Young PSR
PSR 0540-69	50 ms	4.79×10^{-13}	1.5×10^{38}	1.67	5.0×10^{12}	1.81	50	Young PSR
J0205+6449	65 ms	1.93×10^{-13}	2.7×10^{37}	5.4	3.6×10^{12}		2.6	In 3C 58
PSR 1509-58	150 ms	1.54×10^{-12}	1.8×10^{37}	1.55	1.5×10^{13}	2.84	4.2	Young PSR
Geminga	237 ms	1.14×10^{-14}	3.4×10^{34}	330	3.3×10^{12}		0.16	gamma PSR
J1852+0040	105 ms	9×10^{-18}	3×10^{32}	2×10^5	3×10^{10}		7	CCO in Kes 79
1E 1207.4-5209	424 ms	1.4×10^{-14}	7×10^{33}	480	2.5×10^{12}		2.2	CCO?
1E 2259+586	6.98 s	6.2×10^{-13}	7×10^{31}	180	7×10^{13}		4	AXP
SGR 0526-66	8 s						50	SGR
J1747-2958	99 ms	6.15×10^{-14}	2.5×10^{36}	25	2.5×10^{12}		5	High-vel. PSR

The braking index, n , is a measure of how well the slowing pulsar follows this strictly electromagnetic behaviour. Some pulsars are steady enough so the period can be well measured for a long time and it is possible to determine \ddot{P} , the rate of change of \dot{P} . This enables us to calculate the braking index since

$$n = \Omega \ddot{\Omega} / \dot{\Omega}^2 = 2 - P \ddot{P} / \dot{P}^2.$$

If the radiation from the pulsar is only electromagnetic, then $n = 3.00$. Table 9.1 gives the value of n as measured from a few young pulsars. Values < 3 indicate that some of the torque is from particles accelerated in the magnetosphere and the source of the relativistic wind.

All things considered, the dipole model works very well for the Crab pulsar but not for the PSR/SNR pairs 3C 58 and G11.2-0.3. Here the characteristic age is much larger than the age of the remnant.

Figure 9.5 shows the P - \dot{P} diagram which summarises the observed characteristics of more than 1400 pulsars. Particular values for \dot{E} , A and B are indicated by diagonal lines. The mass of points at the centre are the radio pulsars, most with periods between 0.1 and 3 s. The millisecond pulsars are at the lower left. These have been members of binary systems and have been spun up by accretion of

material from their companions. These are discussed in Chapter 11. X-ray-emitting pulsars generally have a high value of \dot{E} . The anomalous X-ray pulsars, which are discussed in Section 9.10, are clustered at the top right.

9.4 | The glitch

An extremely regular period, increasing steadily with time, is the hallmark of a pulsar. Some, however, show discontinuities or ‘glitches’ in the timing. The radio signal from the Vela pulsar is monitored daily and a glitch is found to occur roughly every 2.5 years (Fig. 9.6). During a glitch, the frequency of radio pulsations, ν , increases suddenly by ≈ 1 part in 10^6 , and the frequency derivative, $\dot{\nu}$, increases by ≈ 1 part in 10^2 and then recovers steadily from this transient change. Duration of the recovery is 10–100 days. Imagine a star quake changing the shape of the crust and suddenly lowering the moment of inertia by a small amount. Conservation of angular momentum requires a small jump in spin rate. The crust then relaxes back to a new equilibrium configuration.

Glitches offer a rare opportunity to learn something about the interior of the neutron star, and much has been learned from radio-timing data alone (Link *et al.*, 1992). A glitch, which can have

Fig 9.5 The $P - \dot{P}$ diagram for 1474 pulsars, each indicated by a point. AXPs are additionally marked by a triangle, and pulsars in binary systems are marked by a box. Circles show four pulsars in supernova remnants mentioned in this chapter. The SNR are, left to right; the Crab Nebula, SNR 0540-69.3 in the LMC, 3C 58, and the Vela SNR. Straight lines show constant values of spindown energy, E (ergs s^{-1}); magnetic field, B (G); and age, A (years).

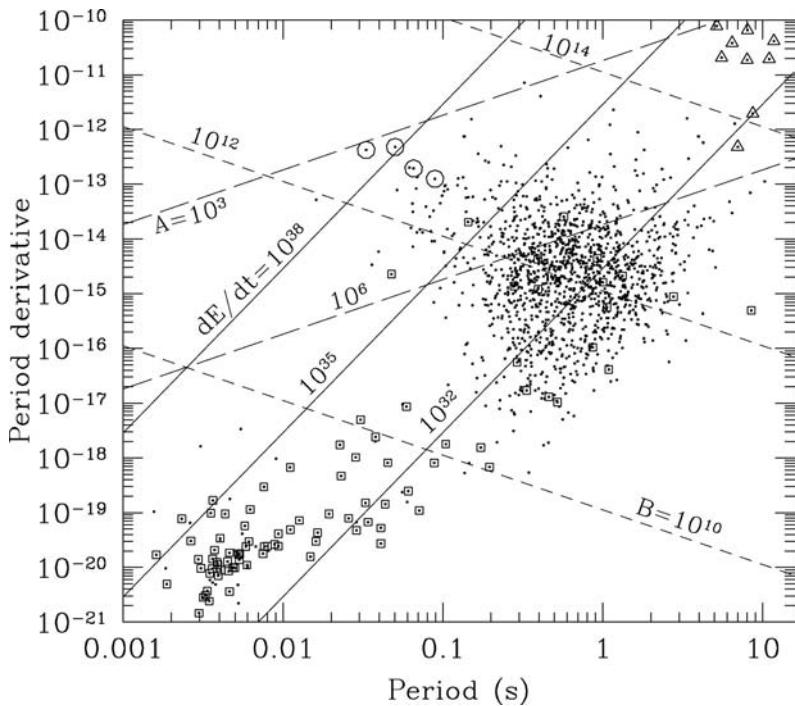
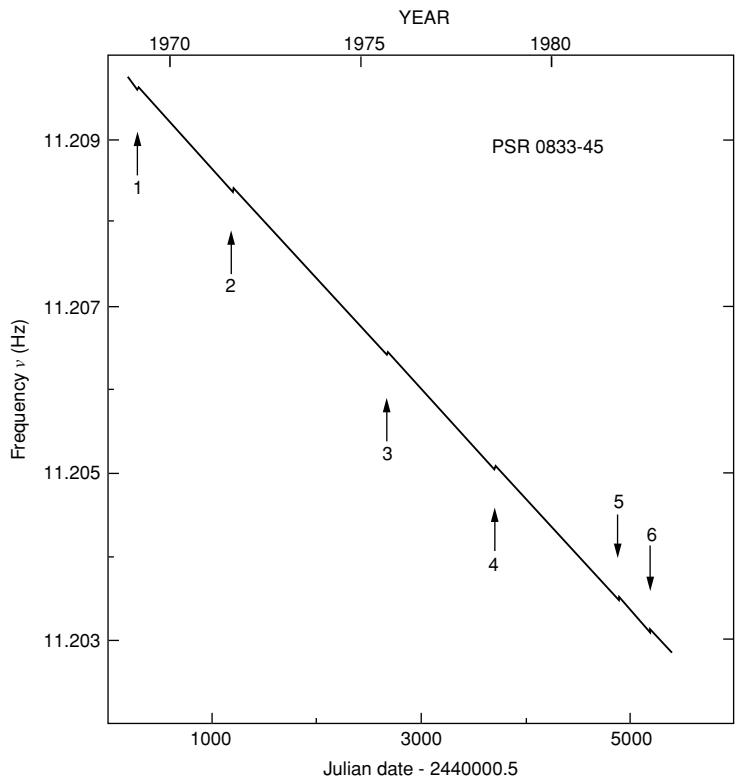


Fig 9.6 Observed Vela pulsar glitches. This plot shows measured pulsar frequency vs time. During the time span of 15 years the frequency decrease is only a change of 0.05 per cent. Glitches are indicated by vertical arrows. Major glitches are numbered and on average occur every 2.5 years (from Cordes et al., 1988).



a duration <1 min, indicates a sudden change in configuration of the solid crust and/or a change in the coupling of the crust, which rotates at the observed rate, to the loosely coupled stellar interior. Some part of the interior, not tied strongly to the torque-producing magnetic field, gets a little ahead of the crust, then suddenly locks to and speeds the rotation of the crust. Current theory favours superfluid protons in the inner crust as the loose component. Whatever the mechanism, during the glitch, rotational energy should be dissipated in the interior and will heat the region where the change occurred. The timescale of energy flow to the surface depends on the size of the star, the nature of the glitch mechanism and the structure of the outer layers. The amount of thermal energy generated and conducted to the surface could be enough to cause an observable increase in temperature of the surface (van Riper *et al.*, 1991). Attempts were made to observe a change in the blackbody X-ray spectrum of the Vela pulsar soon after the glitches of October 1996 and January 2000 (Seward *et al.*, 2000, Helfand *et al.*, 2001). No changes in pulse shape or spectrum were detected. It therefore seems unlikely that useful information about glitches will be obtained from X-ray observations of the surface temperature.

9.5 Pulsed radiation

Most ‘young’ (low P , high \dot{E}) pulsars emit pulsed X-rays, and spectra indicate both thermal and non-thermal origins. Thermal X-rays originate at the surface, non-thermal in the magnetosphere. The Crab pulsar is a magnificent example of non-thermal pulsed radiation. There are two pulses per cycle at all energies from radio to γ rays, and pulse shape and phase do not vary much between the different wavebands (Fig. 9.2). For this object, all pulses probably originate in the same region of the magnetosphere. As a matter of interest, the pulsed radiation represents only ≈ 1 per cent of the Crab pulsar spindown energy.

Theories concerning the production of non-thermal pulses are reviewed by Kaspi *et al.* (2006). Briefly, a rotating magnetised neutron star has a strong surface charge. Electric fields pull charge

from the surface to fill sections of the magnetosphere. The magnetosphere is filled with regions of positive and negative charge divided by the ‘null-charge surface’ where the magnetic field is perpendicular to Ω (Fig. 9.1). The magnetic field and space charge co-rotate with the neutron star out to a radius where the co-rotation velocity = c . Co-rotation must stop here, far from the star. (This radius is 1600 km for the Crab pulsar.) Traditionally, this ‘speed-of-light cylinder’ is thought to be the origin of emission of the radio pulses. Thus a pulse of radiation is observed when one of the magnetic poles rotates through the observer’s line of sight. Radio waves are produced by the synchrotron process when the electrons stop co-rotating and move across magnetic field lines. The width of the resultant beam of radiation is unknown. If it is narrow, like a lighthouse beam, there follows the interesting consequence that most radio pulsars are unobservable because we happen not to be illuminated by the rotating beams.

Two regions have been proposed as the origin of high-energy pulses. Polar-cap models assume pair production in the magnetosphere at the poles, close to the surface, and photon generation by the inverse-Compton process or by curvature radiation (see Chapter 2).

Some of these particles travel back to the surface, and polar-cap heating is a feature of these models. Other models assume the origin of pulsed emission is the ‘outer gap’, a region between the last open field lines and the null-charge surface (see Fig. 9.1). The origin of the radiating particles is a pair cascade, and photons come from curvature radiation or the inverse-Compton process. Again some heating is expected from energetic particles returning to the surface. Attempts to use these models to predict spectra and pulse shape have been only partially successful. Irrespective of mechanism, it is remarkable that some pulsars can be observed in the energy range 100–400 MeV. Indeed, most of the pulsed energy of the Vela pulsar and Geminga is in the γ -ray band, and these objects can be thought of as γ -ray pulsars. The Vela pulsar radiates ~ 10 per cent of \dot{E} as pulsed γ rays.

Thermal radiation is better understood. The interior of a young neutron star is expected to

be hot and is surrounded by the insulating crust. Since heat is conducted more easily along, rather than across, field lines, the magnetic poles should be hotter than the rest of the surface. Temperatures in the range 10^5 – 10^6 K are expected, and the blackbody spectrum peaks in or close to the X-ray band. As just mentioned, surface heating is also expected from non-thermal processes.

Interpretation of any observed thermal pulses may also be complicated by general relativity. The strong gravitational field above the surface may bend rays from the far side of the star towards the observer. This effect tends to wash out pulse structure. There are also spectral complications from the neutron star's atmosphere, which are discussed under cooling (Section 9.7).

9.6 | Structure of neutron stars

Using our knowledge of nuclear forces and the structure of atomic nuclei, the structure of a neutron star has been predicted. The interior structure is described by the relationship of volume, pressure and temperature – an equation of state. The inner part of the star is a core consisting of a degenerate neutron fluid mixed with a few per cent protons and electrons. Density ranges from 10^{14} to 10^{15} g cm $^{-3}$, and the core radius is about 10 km. The liquid neutron core is thought to be a superfluid where viscosity is zero and angular momentum is quantised in vortices. Magnetic field lines thread the interior through these vortices. Colour Plate 24 illustrates some of these concepts. The centre, where density is greatest, is unknown territory. It may consist of ‘exotic’ material with mesons and other elementary particles. Surrounding the core is a crust of thickness ≈ 1 km consisting of ions and electrons, changing to neutron-rich nuclei, neutrons and electrons in the inner layers. Coulomb forces between the ions make the crust rigid, but protons in the inner crust can be superfluid. The outermost part is a gaseous atmosphere which was originally Fe, but since the star is expected to accrete matter from ejected material and the ISM, there should be lighter elements, too. The atmosphere will be ~ 1 cm thick and strongly stratified. Any X-rays emitted will originate in only the lightest elements.

Central densities predicted by different equations of state range from 10^{14} to 10^{15} g/cm 3 . At the higher densities, some theories predict the existence of exotic particles (mesons, quarks) in the core. Predictions are probably secure for the crust and outer core, but when the density approaches 10^{15} g cm $^{-3}$, the behaviour of nuclear material is unknown. A stiff equation of state predicts a radius of 15 km, whereas a soft equation of state predicts a radius of 7 km. If the mass is greater than $\sim 3 M_\odot$, then a black hole is formed. If the mass is less than $\sim 0.1 M_\odot$, the density is low enough so that neutrons are unstable to beta decay, and the assembly will disintegrate.

Present theory favours a radius of 12–15 km and maximum mass of $2 M_\odot$ (Lattimer & Prakash, 2004). Observations of binary systems yield mass measurements for several neutron stars (see Chapter 11). Their precision varies, but the few well-determined measures from binary radio pulsars are all between 1.3 and $1.4 M_\odot$, exactly the mass expected if neutron stars are formed by gravitational collapse of white dwarf material at the Chandrasekhar limit of $1.44 M_\odot$.

So we have here a bizarre object built of material with densities far outside our normal range of experience. The star is shaped by gravity and (as we shall see) magnetism, and the fields and forces are titanic. Because many neutron stars radiate strongly in the X-ray band, the X-rays may provide answers about structure and composition.

9.7 | Cooling

A neutron star, born in gravitational collapse with initial temperature $\sim 10^{11}$ K, cools rapidly through neutrino emission from the core. At an age of 10–100 years, the crust is thermally coupled to the core and is transparent to neutrinos. The surface temperature, T , is 10^6 – 10^7 K. Core neutrino emission dominates the cooling until an age of $\sim 10^5$ years, after which photon emission from the surface cools the star.

At an age of 100 years, observations can separate a new neutron star from the cloud of expanding debris. It should appear as a prominent X-ray source since, at $T \sim 10^6$ K, much of the blackbody spectrum is in the X-ray band. The spectral shape

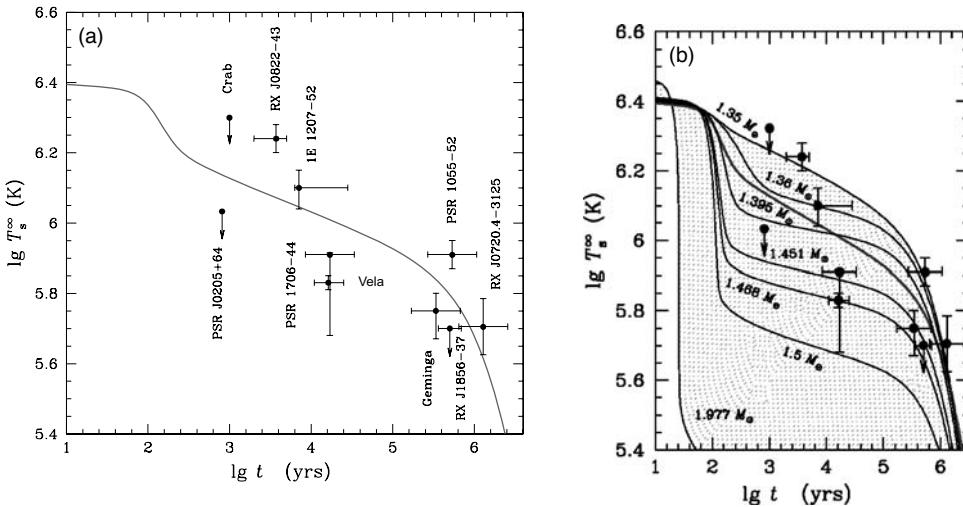


Fig 9.7 (a) Temperature of neutron star surface as a function of age (log-log plots). The calculation for a $1.3 M_\odot$ neutron star is compared with measured blackbody temperatures of several pulsars. (b) Figure 9.7a overlaid with results from a model which allows for proton superfluidity in the crust. This shows how the direct Urca process cools the star rapidly at higher masses. Curves are for several masses ranging from a $1.33 M_\odot$ star to a $2.00 M_\odot$ star (from Yakovlev & Pethick, 2004).

will yield a value for T . Even if T is low and only the tail of the blackbody spectrum is measured, T can be estimated from $L_X = 4\pi R^2 \sigma T^4$ using a reasonable value for R .

Figure 9.7 shows calculations of neutron star cooling curves (surface temperature as a function of time). The cooling curve in Fig. 9.7a is for a star with mass $1.3 M_\odot$. This curve is applicable to neutron stars having this mass or less, independent of the equation of state (soft or stiff) used. The figure also shows temperatures measured for 10 neutron stars. There is considerable scatter.

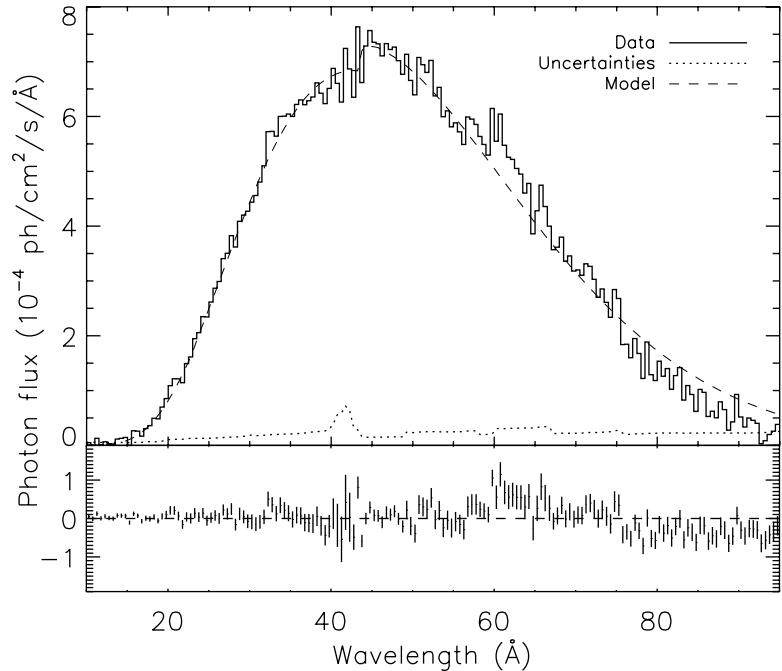
For masses of $1.3 M_\odot$ or less, neutrinos (ν) are generated by the modified Urca process, $n \rightarrow p + e^- + \nu$, then $e^- + p \rightarrow n + \nu$. This Urca process is rather slow because a third nearby particle is necessary to conserve both energy and momentum. For higher-neutron star masses, with consequent higher core densities and higher fractions of protons in the core, a direct Urca process with no spectator particle is possible, the rate of neutrino emission increases and cooling is much faster. Figure 9.7b uses a stiff equation of state and assumes that some of the protons are superconducting. This superconductivity both reduces the rate of the Urca process and broadens the transition to the direct Urca process. The standard curve

of Fig. 9.7a is shown as a heavy line. Note that at an age of 10^4 years, low-mass stars are warmer than predicted by the standard curve and that high-mass stars cool more rapidly. At $2.0 M_\odot$, the star has cooled in <100 years. By choosing the mass, this theory can be made to fit any data point falling in the shaded area. By choosing individual masses between 1.35 and $1.47 M_\odot$, all the data shown can be fit by this process.

Rapid neutrino cooling will also occur in the presence of exotic particles such as mesons in the core. Since the central density is probably beyond reasonable extrapolation of known nuclear forces, the existence of mesons or quarks in the core is possible. Note that the only measures of neutron star mass have been in binary systems (Chapter 11) and, although a larger range is possible, all observations can be fit with masses in the range 1.3 – $1.4 M_\odot$. One may therefore not be free to use all the curves shown in Fig. 9.7.

Initial hopes that surface temperatures could be determined from measured spectra in a straightforward way have not been realised. The neutron star atmosphere is easily thick enough to absorb X-rays, and high-energy photons come from deeper layers than photons of lower energy. Observed spectra, fit directly with blackbody

Fig 9.8 A 500-ks (6 days) Chandra grating spectrum of the nearby neutron star RX J185635-3754. Instrument efficiency has been removed to show the photon spectrum. Dashed line shows a 61-eV blackbody spectrum which is a good fit. The lower curve shows deviations of data from this curve (from Drake *et al.*, 2002).



shapes, yield temperatures sometimes a factor of 2 higher than the actual surface temperature. Strong magnetic fields can also distort spectra.

One of the nearest neutron stars is RX J185635-3754, discovered by Fred Walter *et al.* (1996) in a ROSAT observation. The identification is based on the thermal spectrum with temperature of 57 eV and the low luminosity of the extremely faint (25.7 mag) optical counterpart. Distance has been measured by parallax to be 117 pc. A long Chandra observation using a grating was made to determine the temperature and to search for spectral features. If emission or absorption features could be identified, the composition of the atmosphere might be known. There should also be a large gravitational red shift, and for a given mass, the radius would be established. However, the observed spectrum, shown in Fig. 9.8, has no features and is well fit by a blackbody spectrum with temperature 7×10^5 K ($kT = 61$ eV). The implied radius is 4–8 km, rather small to fit any of the current models. This turns out to be the case for all observations of neutron star surface temperatures. Radii determined from measured L_X or T assuming a blackbody spectrum are too small. Introduction of

a model atmosphere gives radii closer to expectations but introduces uncertainties in atmosphere composition and magnetic field effects.

Object 1E 1207.4-5209, an object at the centre of the supernova remnant PKS 1209-52, is different. Figure 9.9 shows the X-ray spectrum measured by XMM. There are significant deviations from a blackbody continuum. Two strong and two weak minima are interpreted as absorption bands caused by cyclotron scattering of electrons in the magnetosphere. Since 10 per cent pulsations with period 0.424 s have been observed, the object is known to have a magnetic field, and the value derived from the cyclotron resonance energies is $\sim 8 \times 10^{10}$ G. This is an order of magnitude smaller than the field derived from P and \dot{P} and listed in Table 9.2. The continuum curve in Fig. 9.9 is made of two blackbody spectra with $kT = 0.21$ and 0.40 keV, corresponding to radii of 3.0 and 0.25 km. As usual, the blackbody temperatures are probably too high. The presence of a magnetic field apparently makes the situation more complex than that of a single hot spot on the surface.

Since the age of a neutron star in a supernova remnant is sometimes well determined through the historical record (see Chapter 8), these objects

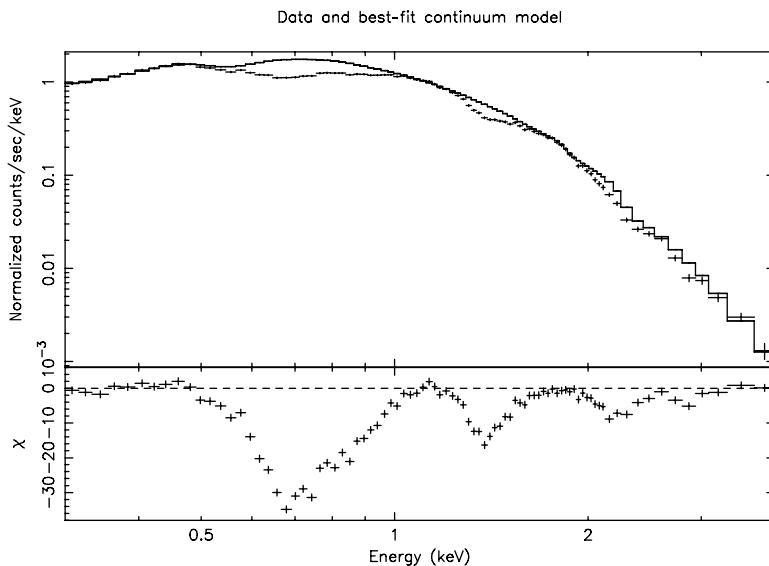


Fig 9.9 Spectrum of the isolated neutron star IE 1207.4-5209 as recorded in a 54-hour XMM observation. A best-fit theoretical continuum is also plotted. The lower curve shows deviation of the data from this continuum. The minima are interpreted as absorption of X-rays by electrons spiraling around magnetic field lines (from Bignami et al., 2004).

make good data points for comparison with cooling calculations. Modelling of the atmosphere is needed, and there are many processes and parameters that enter into the cooling calculations. We can now show that some data are consistent with some calculations. It is difficult to determine the internal parameters, but this may eventually be possible for at least some neutron stars. A good review of the properties of neutron stars can be found in Lattimer and Prakash (2004).

9.8 Pulsar wind nebulae

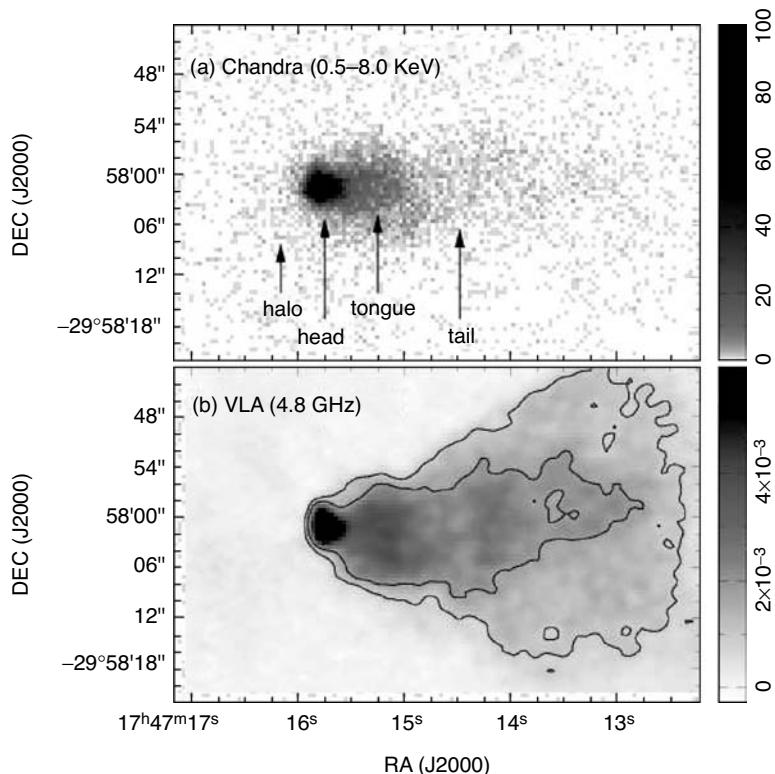
Your attention is now directed from the neutron star itself to the PWN, the dominant feature in most X-ray observations. The Crab pulsar is surrounded by a PWN which radiates strongly from the radio band all the way up to TeV γ rays. Ninety-five per cent of the X-ray emission shown in Fig. 9.4 is from this PWN. All fast pulsars with strong magnetic fields are surrounded by such PWN. Even if the pulsar itself is invisible because of beaming, the PWN, which radiates isotropically, can easily be detected. The remnants MSH 11-62 and MSH 15-56 are examples.

A PWN is a most unusual state of matter. It consists of a cloud of highly relativistic particles,

both electrons and positrons, and a magnetic field of strength 10^{-5} – 10^{-3} G. (The average interstellar field is 5×10^{-6} G.) We see radiation from electrons moving in the magnetic field. The resultant synchrotron spectrum is a power-law continuum, $N = k(h\nu)^{-\alpha}$, and this is accepted as the signature of non-thermal emission. For the Crab, $\alpha \approx 0.25$ in the radio band and ≈ 2.0 from optical through X- to γ rays. Note the spectral break in Fig. 9.3, where the power-law index changes value in the infrared.

The dynamics of PWNe were first set forth in one-dimensional models (Pacini & Salvati, 1973; Rees & Gunn, 1974; Kennel & Coroniti, 1984) which were applied to the Crab. In these models, particles and field flow from the pulsar at the speed of light and, by the time they reach the filamentary envelope, have slowed to the observed expansion velocity of ~ 1500 km s $^{-1}$. To slow down, the particles must pass through a shock which should be located close to the pulsar, where emission is brightest. At the shock, particles are accelerated into the power-law spectrum. Electron velocities, all directed outwards before the shock, are oriented randomly after the shock. The electron speed remains $\approx c$ throughout the nebula. The appearance of the PWN should be that of a hollow feature close to the pulsar since there is no radiation from the orderly flow before the

Fig 9.10 Radio and X-ray emission from the PWN associated with the high-velocity PSR J1747-2958. The pulsar is traveling east (right to left). Ram pressure of the ISM causes the PWN to form behind the pulsar (from Gaensler et al., 2004).



shock, then a bright shock around the pulsar, then a decrease in brightness as the particles lose energy travelling outward to the envelope. Because the field of the spinning pulsar has cylindrical geometry, the shock appears as a torus surrounding the pulsar. Since radiation from a relativistic electron is strongly concentrated in the direction of motion, Doppler brightening makes the near side of the torus appear brighter. This geometry can be seen in Fig. 9.4. Particles lose energy during the journey from shock to envelope, and this is reflected in the size of the PWN. The Crab PWN appears to fill the nebula when observed in the radio band and gets smaller as photon energy increases. Colour Plate 23 shows the X-ray PWN only half the size of the optical nebula. The synchrotron lifetime of an electron of energy E in a magnetic field B is $\propto E^{-1}B^{-2}$. Electrons which radiate X-rays have lifetimes of only a few years and lose their energy before reaching the edge of the nebula. They must be replenished continuously. The strength of the high-energy PWN thus shows the current energy input from the

pulsar. Electrons which radiate in the radio band have lifetimes longer than the age of the nebula, and the luminosity of the radio PWN thus indicates the integrated output of the pulsar. It is usually assumed that energy in the PWN is equally divided between particles and magnetic field, but it is not clear that this is necessary.

The morphology of the inner Crab PWN – a torus close to the pulsar and jet(s) along the rotation axis – is also seen around other young pulsars (Fig. 9.12 and Colour Plate 26) Doppler beaming usually causes only the jet coming towards us to be visible. Note that observation of a jet seems to give the orientation of the spin axis. In cases where proper motion has been measured, such as the Vela pulsar, the velocity vector is found to be close to the spin axis.

Because synchrotron luminosity is proportional to B^2 , a bright PWN implies strong B , and PWNe may be characterised by an average value of B therein. Accounting for radiation at all energies (not just X-rays), some values are listed in Table 9.2.

Table 9.2 \bar{B} for several PWN.	
Crab	$\approx 6 \times 10^{-4}$ G
3C 58	$\approx 3 \times 10^{-5}$ G
Vela SNR	$\approx 1.5 \times 10^{-4}$ G

The PWN luminosity is found to depend strongly on \dot{E} . For the Einstein band, 0.2–4 keV, including radiation from both pulsar and PWN, $L_X \approx 2.5 \times 10^{-17} \dot{E}^{1.39}$ (Seward & Wang, 1988). For all cases where they were resolved, the L_X of the PWN was considerably more than that of the pulsar itself. For the pointlike pulsar, in the ROSAT band, 0.1–2.5 keV, $L_X \approx 10^{-3} \dot{E}$ (Becker & Trümper, 1997). These two relations do not agree for $\dot{E} < 10^{34}$ ergs s $^{-1}$, and it has been proposed (Gotthelf, 2004) that PWN are not formed below a threshold of $\dot{E} = 4 \times 10^{34}$ ergs s $^{-1}$. Since observation of very faint extended emission above the background level is quite difficult, this is currently an open question.

As the surrounding supernova remnant ages, an internal PWN evolves. The pulsar is continuously pumping energy into the PWN, so it pushes into the surrounding hot gas and expands. When the reverse shock from the ejecta/ISM interaction reaches the expanding PWN, compression and some instability of the boundary are expected. At late times, since some pulsars are born with high kick velocities (up to 1000 km s $^{-1}$), the pulsar can travel outside the PWN envelope, leaving a relic PWN and creating a new PWN along its path (for a review, see Gaensler & Slane, 2006). An example of this is the remnant N206 in the LMC, which has a compact X-ray source at the edge with trail extending back towards the centre of the remnant (Williams *et al.*, 2005).

Cometary PWNe have also been observed behind fast pulsars travelling through the ISM. X-ray and radio emission associated with the young pulsar J1747-2958 (Fig. 9.10) have a bow-shock morphology caused by motion of the pulsar through the ISM. The radio morphology of this PWN, known as G359.23-0.82, or more colourfully as the ‘Mouse’, stretches 12' along the direction of motion. Analysis shows the pulsar to be moving at 600 km s $^{-1}$ through a medium with density of 0.3 atoms cm $^{-3}$. The X-ray PWN is also

stretched along the path of the pulsar. In Fig. 9.10, the head is in part the unresolved pulsar and in part a broader component. The tongue is the wind termination shock and the tail the postshock particles. There are, at this time, four other pulsars which show cometary morphology.

9.9 | Central compact objects

The *central compact objects* (CCOs) are pointlike X-ray sources found within a few supernova remnants. Because there is no surrounding or nearby PWN, they cannot be energetic pulsars. Radio searches have not detected them, so they are sometimes called ‘radio quiet’. Pulsations have been detected from two: 1E 1207.4-5209 in the remnant PKS 1209-51/52 ($P = 0.424$ s) and PSR J1852+0040 in the remnant Kes 79 ($P = 0.105$ s). Their spindown rates (\dot{P}) are low and give characteristic ages much greater than the estimated ages of the remnants. Spectra are close to blackbody with $kT \approx 0.4$ keV and $L_X = 10^{33} - 10^{34}$ ergs s $^{-1}$. There are CCOs in the remnants Cas A, Pup A, G266.2-1.2 and G347.3-0.5 from which no pulsed component has yet been found. The upper limit on pulsed X-rays from the central source in Cas A is 12 per cent. The nature of these objects is not well understood. Cooling neutron stars is the obvious choice, but the spectra and the generally weak pulsations are yet to be reconciled with this picture.

9.10 | Anomalous pulsars, soft-gamma repeaters and magnetars

The AXPs and SGRs occupy the upper right corner of the $P-\dot{P}$ diagram (Fig. 9.5). They are all quite luminous with $L_X \approx 10^{36}$ ergs s $^{-1}$, an order of magnitude greater than \dot{E} . So unlike all of the rest of the ~ 1400 objects plotted, these seven need an additional source of energy. Accretion was the obvious choice, and the first two AXP discovered, 1E 2259+586 (Fahlman & Gregory, 1981) and 1E 1048.1-5937 (Seward *et al.*, 1986), were thought for years to be members of binary systems. Diligent

searches for companions or evidence of orbital motion, however, found nothing, so these were called ‘anomalous’.

The first SGR discovered, SGR 0525-69, was found by the γ -burst detection network and was located to be within the supernova remnant N49 in the LMC (discussion and figure in Chapter 17). A subsequent series of pulses, the repeating characteristic of the class, was observed soon after. It took a while to realise that the SGR were different from the extragalactic γ -burst sources and that this object was really associated with the remnant. A ROSAT observation was the first to verify an unresolved source at the SGR location (Marsden *et al.*, 1996). Fifteen AXPs/SGR are now known (nine are indicated in Fig. 9.5), and the discovery of γ -ray bursts from AXP 1E 2259+586 showed that there is little difference between these two groups.

As a class, the AXPs/SGR appear to have the following properties:

- Their X-ray emission is periodic with observed periods 5–12 s, long for young neutron stars.
- The observed \dot{P} is high, implying $\dot{E} \approx 10^{33}\text{--}10^{34}$ ergs s^{-1} , much too low to account for the observed steady luminosities, $L_X \approx 10^{35}\text{--}10^{36}$ ergs s^{-1} . The calculated magnetic field, B , is enormous: $10^{14}\text{--}10^{15}$ G at the poles. Thompson and Duncan (1995) proposed that the energy source is the strong magnetic field. They suggested a mechanism whereby energy from the field might heat the crust of the neutron star to account for the observed luminosity and occasionally crack the crust to release bursts of high-energy radiation. They named these sources *magnetars*.
- At unpredictable intervals, there are soft γ -ray/hard X-ray bursts with energy up to $\approx 10^{41}$ ergs/burst.
- There are rare giant bursts. The 5 March 1979 γ -ray burst from SGR 0525-69 lasted 0.2 s, and the energy in the burst was 5×10^{44} ergs. The observed afterglow lasted 3 min and carried a clear 8-s periodic signal. On 22 August 1998, SGR 1900+14 produced a 1-s outburst containing 2×10^{43} ergs with a 6-min tail and 5.2-s period. The largest outburst yet recorded occurred

27 December 2004. A 0.2-s outburst from SGR 1806-20 had a peak luminosity of 10^{47} ergs s^{-1} and bolometric energy of 4×10^{46} ergs. This produced a detectable disturbance in the Earth’s ionosphere, and the 6-min tail showed 7.6-s pulsations.

- The quiescent X-ray emission is highly variable. Although the sources we know are usually bright, a new transient AXP, XTE J1810-197, was discovered with XTE when it brightened suddenly from a state of much lower luminosity (Ibrahim *et al.*, 2004). Searches of data archives revealed a faint source at this location with approximately constant flux for an interval of ~ 20 years. The outburst produced a factor of 100 increase in luminosity. A period of 5.5 s and rapid spindown were observed, characteristic of an AXP. The outburst faded with an *e*-folding time of ~ 300 days (Ibrahim *et al.*, 2004).

Half the known magnetars are associated with supernova remnants and must have ages of only a few thousand years. The others, if born in a remnant, have outlasted all signs of it or have travelled far from their point of origin. The observation of a transient AXP implies that there may be a large population of faint magnetars not yet identified.

9.11 Some supernova remnants containing a variety of neutron stars

Many remnants contain neutron stars, and both remnant and central object are manifested in different ways. Table 9.3 lists some well-known examples, most of which are described in this chapter.

9.11.1 The Vela supernova remnant

The Vela supernova remnant (or Vela XYZ, a complex of three radio sources) is only 280 pc distant. The optical SN must have been spectacular. SN 1006 at magnitude –9 was as bright as the quarter Moon. Even allowing that a type II is not as bright as a type I, the nearer Vela SN would have been 10 times brighter. The age of the remnant

Table 9.3 Selected supernova remnants and associated neutron stars.

Remnant	L_X total (ergs s $^{-1}$)	L_X NS (ergs s $^{-1}$)	L_X PWN (ergs s $^{-1}$)	Distance (kpc)	Angular size (arcmin)	Size (pc)	Comment
Crab	3×10^{37}	1×10^{36}	3×10^{37}	2	5 × 7	2.9 × 4.1	Radio PSR
0540-69.3	2.1×10^{37}	4×10^{36}	1.3×10^{37}	50	1.1 × 1.2	16 × 17	Radio PSR
Vela SNR	3×10^{35}	2×10^{32}	5×10^{32}	0.29	480 dia	40 dia	Radio PSR
3C 58	3×10^{34}	3×10^{33}	2×10^{34}	3.2	6.5 × 5	6.0 × 4.6	Radio PSR
Cas A	1×10^{37}	2×10^{33}	–	2.6	4 dia	3.0 dia	CCO
Kes 79	2.5×10^{36}	3×10^{33}	–	7	11 dia	22 dia	CCO
CTB 109	4×10^{37}	1.5×10^{36}	–	4	25 × 32	29 × 36	AXP
Kes 73	9×10^{36}	4×10^{35}	–	7	4.6 dia	9.3 dia	AXP
N 49	2.2×10^{37}	1.2×10^{36}	–	50	1.4 dia	20 dia	SGR
The Mouse	5×10^{34}	3×10^{34}	2×10^{34}	5	0.1 × 0.4	0.15 × 0.6	High-vel. PSR

is estimated as 10 000 years. Surely this SN must have made a strong impression on any people living south of latitude $\approx 10^\circ$ (precession is important). There is even a book proposing a connection between this event and the dawn of civilisation (Michanowsky, 1977).

X-rays from the Vela remnant were first detected in May 1968 as a very soft source in the southern sky (Grader *et al.*, 1970). This source was identified as a superposition of both the Vela

remnant and the remnant Puppis A by Palmieri *et al.* (1971). Although maps of the emission were obtained with rocket-borne detectors and with the SAS-3 satellite, not until Einstein was the structure of the emission from this region clearly visible. The most detailed map so far was made by ROSAT (Lu & Aschenbach, 2000) and is shown in Fig. 9.11. The outer shock is visible in a few places but is not bright in X-rays. The brightest X-rays come from great irregular patches of gas inside the boundary,

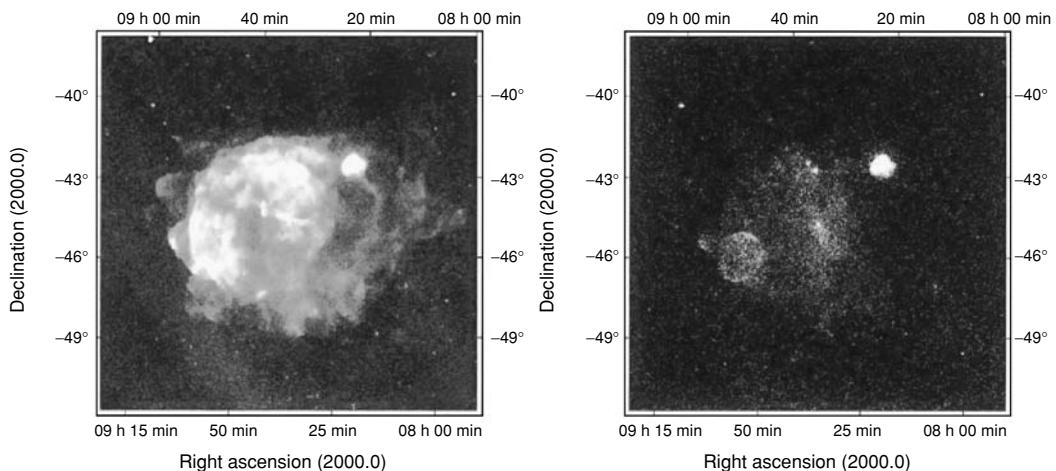


Fig 9.11 The Vela supernova remnant as mapped by the ROSAT all-sky survey in two energy bands: at left 0.1–2.4 keV; at right, 1.3–2.4 keV. This remnant is large at 8° in diameter. The smaller, more distant remnant Puppis A shines through the northwest edge. The effective exposure for each element of this map is only 11 min. These remnants are the brightest soft X-ray sources in the sky (the Sun excepted; see Fig. 1.6). A third remnant lies behind the southeast quadrant but, at the left, is masked by soft emission from the Vela remnant. The picture at the right, with soft emission excluded, shows the shell-like Vela Jr, the Vela pulsar/PWN at the centre of the Vela remnant, and Pup A (from Aschenbach 1998).

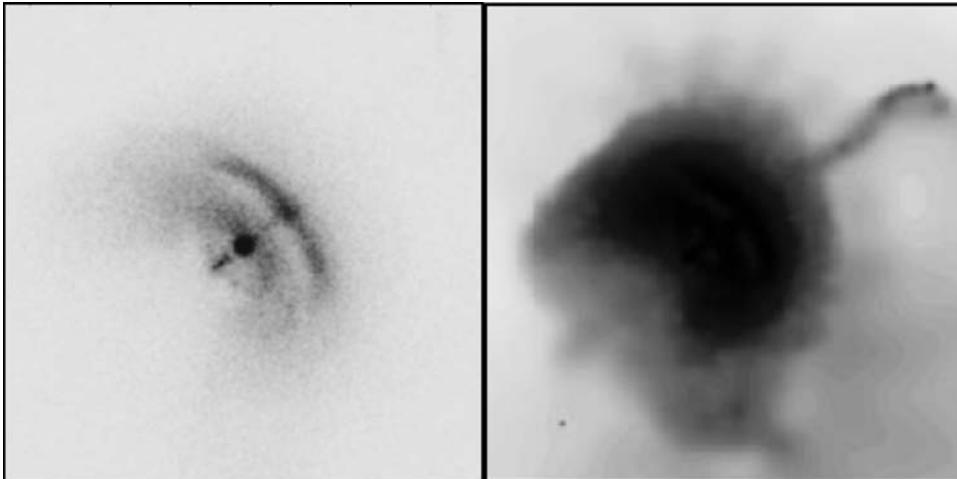


Fig 9.12 Chandra images of the Vela pulsar and its PWN. (left) Two 13-hour HRC exposures were added to make this image, which is 2.0 arcmin on a side. The pulsar is the bright point source. The torus seems split into two parts. Jets propagate in both directions along the pulsar spin axis. Note that the extent of this nebula is only $\approx 1'$. (right) Eight 6-hour ACIS pointings taken over a 9-month interval have been summed to produce this picture of the faint northern jet. The centre of the PWN is overexposed (from Helfand et al., 2001; Pavlov et al., 2003).

and these areas are also filled with bright optical filaments. Apparently the SN explosion occurred before the stellar wind could remove lumpy interstellar material from the near vicinity. The denser regions now show as bright optical filaments surrounded by hot X-ray-emitting gas. Some of this gas is probably that which evaporated from the filament-forming clouds. The X-ray emission is quite soft, satisfying our expectation that older remnants should be cooler.

There are several ‘bullets’ associated with the Vela remnant. These are relatively compact, high-velocity blobs of material which have now travelled farther than the outer shock. The brightest ones and associated bow shocks can be seen in Fig. 9.11 due east and leading the shock by $\sim 1^\circ$. Another surprise from this area came with the discovery by Aschenbach and Lu (2000) of another remnant (in addition to the well-known Pup A) behind the brightest part of the Vela remnant. This remnant, RX 0852.0-4622 (also referred to as Vela Jr), consists of a circular shell with a weak compact object at the centre. It is only visible at energies above ≈ 1 keV.

The pulsar within the Vela remnant was discovered by Large *et al.* (1968) and was the

second pulsar found within a supernova remnant. Its characteristic age of 10^4 years is a good match for the age estimated for the shell of the remnant. This pulsar, 10 times older than the Crab pulsar, was not expected to be as energetic. Its \dot{E} is 66 times less, and in the soft X-ray band, L_X is $\sim 10^4$ times less. The pulsar is surrounded by a PWN which, although weak, is near enough to be well resolved and has interesting peculiarities. Compared to the Crab PWN, the Vela PWN has 10^{-3} the volume and 10^{-4} the X-ray luminosity in the 0.3–10 keV band. Figure 9.12 shows that the bright centre is toroidal like the Crab but appears as two distinct rings. There are also jets along the axis of symmetry, one pointing northwest, one southeast. The faint envelope of the PWN has a bow-shock morphology, and indeed, this is aligned with the direction of motion. The distance (286 ± 16 pc) and velocity (61 ± 2 km s $^{-1}$) have been recently measured using the Long Baseline Array (LBA) radio network by Dodson *et al.* (2004). There are also two outer jets. The brighter points northwest, in the direction of motion, and contains a knot of emission close to the end. This jet also varies in intensity and direction fairly rapidly. Changes occur on a timescale of weeks. Pavlov *et al.* (2003)

attribute these changes to instabilities in a stream of particles with a self-generated magnetic field and propose that the very faint emission seen south of the PWN in Fig. 9.12 is because of particles from the outer jets.

The Vela pulsar/PWN complex shows that the Vela remnant originated in gravitational collapse. In the band 0.2–4 keV, only 0.4 per cent of the emission from the remnant is from the pulsar/PWN, yet the observations show amazing detail. In the X-ray band, only 10 per cent of the pulsar emission is pulsed. The pulsar, however, is a strong source of 100-MeV γ rays. Most of the pulsed energy is in the γ -band, which is quite different than the spectral distribution of energy from the Crab pulsar. Pulsars as well as supernova remnants are individuals.

9.11.2 SNR 0540-69.3

This remnant has been called the ‘twin’ of the Crab Nebula. It contains a bright 50-ms pulsar with characteristic age 1670 years. The pulsar is surrounded by a PWN about the same size as the Crab PWN but only half as luminous. The pulsar itself is actually three times more luminous than the Crab pulsar in X-rays. It was discovered with a series of Einstein observations (Seward *et al.*, 1984) and subsequently found to emit optical and radio pulsations. As can be seen in Colour Plate 25, the PSR/PWN complex is surrounded by a shell of thermal emission, the expected outer shock – the blast wave from the SN explosion. This is of unusual interest because the Crab seems devoid of such a shell. Searches in many wavebands have found nothing outside of the Crab PWN (Frail *et al.*, 1995; Seward *et al.*, 2006). Because SNR 0540-69.3 is in the LMC at a distance of 50 kpc, only Chandra has been able to measure spatial details. But even Chandra cannot resolve features in the $2'' \times 3''$ PWN.

9.11.3 3C 58

This remnant contains a spin-powered pulsar ($P = 0.065$ s) and PWN and is thought to be the remnant of SN 1181. The description of the location in the historical record from China is not accurate enough to identify the remnant without doubt. As

a possible/probable historical remnant, it is worthy of study. The characteristic age of the pulsar is 10^4 years, which contradicts the historical age. Remnant 3C 58 is 2–3 times larger than the Crab, and the pulsar and PWN are a factor of 1000 fainter than the Crab pulsar and PWN. Colour Plate 26 shows a Chandra observation. The PWN which fills the remnant is full of ropey filaments which are not understood. The bright structure close to the pulsar is interpreted as a torus and a westward-pointing jet. There are also faint optical filaments which do not coincide with the PWN structure. There is another remnant, G11.3-0.2, reputed to be from SN 486, which also contains a pulsar with characteristic age 10^4 years. Probably the historical identifications of both remnants are correct and the pulsars have evolved in a way that now gives a characteristic age much greater than the true age. If the identification with SN 1181 is correct, the equations in Section 9.3 can be used to calculate the initial period of the 3C 58 pulsar. In this case, $P_0 = 0.060$ s, only ≈ 10 per cent less than the present period.

9.11.4 CTB 109

This remnant was discovered in an Einstein observation. Gregory and Fahlman (1980) were searching for X-ray emission from the variable radio source GT 2257+585 in Cassiopeia. When the data were processed, in one corner of the field was a large, diffuse source – obviously a supernova remnant but one not in any SNR lists or catalogues. They had discovered a new, rather bright remnant. Although the radio source, CTB 109, was known, it was only 3° from Cas A, the brightest radio source in the sky, and more sensitive radio observations had not been attempted. Not only was this remnant new to our eyes but the structure was unique. A circular half-shell was centred on a bright, unresolved source. A follow-up, deeper observation with the source centred in the field was used to search for pulsations from the central object. The period found was 6.98 s, much longer than for other neutron stars within remnants. Subsequent observations to measure \dot{P} produced irregular results. A decade later, this was recognised as an isolated AXPs. Colour Plate 27 shows an XMM observation of CTB 109.

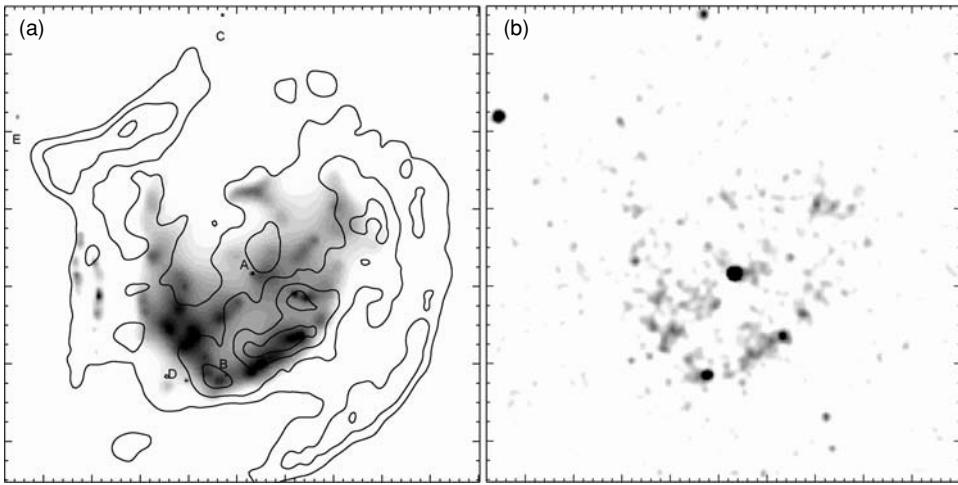


Fig 9.13 The supernova remnant Kes 79 imaged by Chandra in two energy bands. (a) Grey-scale X-ray emission in the 1–8 keV band is overlaid with contours of constant radio surface brightness. The brightest X-ray emission originates from an interior shell and has a thermal spectrum. The unresolved source at the centre labelled ‘A’ is a CCO. Unresolved sources labelled ‘B–E’ are not associated with the remnant. (b) Kes 79 in the 3–5 keV energy band. These data have been smoothed more broadly than the data in Fig. 9.13a. The CCO is more prominent in this band. Both fields are 12 arcmin on a side (from Seward *et al.*, 2003).

The remnant is in contact with a molecular cloud (Tatematsu *et al.*, 1987). The presence of the cloud to the west accounts for the unusual shape. The western half of the remnant has evolved and cooled rapidly in the dense material of the cloud and is now invisible. An extension of the cloud shadows part of the northern remnant and partially accounts for the shape of the feature connecting the central source and shell. There was a flurry of interest in this structure as a possible jet originating at the pulsar, but this feature is now recognised as part of the thermal envelope.

9.11.5 Kes 79

A map of the radio source Kes 79 shows the shell of a middle-aged SNR with fairly strong emission from the interior (Fig. 9.13). The first X-ray observation was by Einstein (Sequoist & Gilmore, 1982), and since the emission seemed to fill the interior, it was thought to contain a possible PWN. A ROSAT observation (to map the PWN) showed weak emission from an outer shell, stronger emission from a shell-like structure in the south central area and no PWN. Chandra was then used to map details of the bright centre and, to our surprise, found an unresolved source at the centre, a weak pointlike

CCO similar to the one in Cas A. There was no surrounding PWN, and the source had a blackbody spectrum with $kT = 0.5$ keV and $L_X \approx 7 \times 10^{33}$ ergs s^{-1} (Seward *et al.*, 2003). An XMM observation revealed strong pulsations with a period of 106 ms. This was unexpected since only one CCO had previously been found to pulse (Gotthelf *et al.*, 2005). Follow-up XMM and Chandra observations were done to measure \dot{P} , and 2 years after the first observation, a period derivative was measured. The pulsar is spinning down very slowly. (See Section 9.3, Table 9.1.) The energy loss is small and the magnetic field is low. This pulsar is thought to have been born with a weak magnetic field, and spindown has been very small. The absence of PWNe indicates that this is probably true for many other CCOs in supernova remnants.

9.12 | Summary

Most isolated neutron stars that have been detected appear as pulsars. They spin with periods ranging from 10^{-3} to 10 s and are determined to have strong magnetic fields. A standard dipole model gives a basis for classification and comparison, although the derived characteristic age is

sometimes much greater than the estimated age of an associated remnant. Those with high rates of rotational energy loss are always accompanied by a diffuse PWN which has a higher X-ray luminosity than the pulsar itself. They are also bright sources of pulsed non-thermal X-rays. These high- \dot{E} systems and the radio pulsars are powered by the loss of rotational energy.

A few objects have relatively slow spin rates and high values of the derived magnetic field. These “magnetars” are believed to be powered by the decay of the strong field.

A few X-ray sources have been observed which show no sign of rotation. Spectra are close to black-body, and these are believed to be cooling neutron stars powered by internal thermal energy. There may be vast numbers of these, and of magnetars, as yet undetected. We have probably seen all the high- \dot{E} pulsars in our Galaxy.

References

Observation

- Abdo, A. A., et al. (2010) Fermi Large Area Telescope observations of the Crab Pulsar and Nebula, *ApJ* **708**, 1254.
- Aschenbach, B. (1998) Discovery of a young nearby supernova remnant, *Nature* **396**, 141.
- Becker, W. & J. Trümper (1997) The X-ray luminosity of rotation-powered neutron stars, *A&A* **326**, 682.
- Bignami, G. F., et al. (2004) 1E 1207.4-5209 – a unique object, *Mem. S. A. It.* **75**, 448.
- Bowyer, S., et al. (1964a) X-ray sources in the Galaxy, *Nature* **201**, 1307.
- Bowyer, S., et al. (1964b) Lunar occultation of X-ray emission from the Crab Nebula, *Science* **146**, 912.
- Brinkmann, W., B. Aschenbach & A. Langmeier (1985) X-ray morphology of the Crab Nebula, *Nature* **313**, 662.
- Cordes, J. M., G. S. Downes & J. Krause-Polstorff (1988) JPL pulsar timing observations. V. Macro- and microjumps in the Vela pulsar 0833-45, *ApJ* **330**, 847.
- Dodson, R., et al. (2004) The Vela pulsar, the key? in *Young Neutron Stars and Their Environments*, IAU Symposium 218, 193, edited by F. Camilo & B. Gaensler, San Francisco: Astronomical Society of the Pacific.
- Drake, J. J., et al. (2002) Is RX J1856.5-3754 a quark star, *ApJ* **572**, 996.
- Fahlman, G. & P. Gregory (1981) An X-ray pulsar in G109.1-1.0, *Nature* **293**, 202.
- Frail, D., et al. (1995) Does the Crab have a shell?, *ApJ* **454**, L129.
- Gaensler, B. M., et al. (2004) The mouse that soared: High-resolution X-ray imaging of the pulsar-powered bow shock G359.23-0.82, *ApJ* **616**, 383.
- Gotthelf, E., (2004) A spin-down power threshold for pulsar wind nebula generation? in *Young Neutron Stars and Their Environments*, IAU Symposium 218, 225, edited by F. Camilo and B. Gaensler, San Francisco: Astronomical Society of the Pacific.
- Gotthelf, E., J. Halpern & F. Seward (2005) Discovery of a 105 ms X-ray pulsar in Kesteven 79: On the nature of compact central objects in supernova remnants, *ApJ* **627**, 390.
- Grader, R. J., et al. (1970) The soft X-ray spectra of three cosmic sources and simultaneous optical observation of Sco XR-1, *ApJ* **159**, 201.
- Gregory, P. & G. Fahlman (1980) New X-ray SNR with compact central source and possible jet, *Nature* **287**, 805.
- Gursky, H., et al. (1963) Further evidence for the existence of galactic X-rays, *Phys Rev Lett* **11**, 5300.
- Harden, F. R., Jr. & F. D. Seward (1984) Einstein observations of the Crab Nebula pulsar, *ApJ* **283**, 279.
- Helfand, D. J., E. V. Gotthelf & J. P. Halpern (2001) The Vela pulsar and its synchrotron nebula, *ApJ* **556**, 380.
- Hester, J. J., et al. (2002) Hubble Space Telescope and Chandra monitoring of the Crab synchrotron nebula, *ApJ Lett* **577**, L49.
- Hewish, A., et al. (1968) Observation of a rapidly pulsating radio source, *Nature* **217**, 709.
- Ibrahim, I. M., et al. (2004) Discovery of a transient magnetar: XTE J1810-197, *ApJ* **609**, L21.
- Large, M. I., A. E. Vaughan & B. Y. Mills (1968) A pulsar supernova association, *Nature* **220**, 340.
- Link, B., R. Epstein & K. van Riper (1992) Pulsar glitches as probes of neutron star interiors, *Nature* **359**, 616.
- Lu, F. J. & B. Aschenbach (2000) Spatially resolved X-ray spectroscopy of the Vela supernova remnant, *A&A* **362**, 1083.
- Marsden, D., et al. (1996) ROSAT observations of the soft gamma-ray burst error box coincident with the supernova remnant N49, *ApJ* **470**, 573.
- NASA/CXC/Slane et al. (2004) http://chandra.harvard.edu/photo/2004/3c58/3c58_xray_3color.jpg.
- NASA/CXC/Seward et al. (2006) <http://chandra.harvard.edu/photo/2008/crab>.
- NASA/ESA/Allison Loll/Jeff Hester (2005) <http://www.spacetelescope.org/images/heic0515a/>.
- Palmieri, T. M., et al. (1971) Soft X-rays from two supernova remnants *ApJ* **164**, 61.

- Pavlov, G., et al. (2003) The variable jet of the Vela pulsar, *ApJ* **591**, 1157.
- Sasaki, M., et al. (2004) XMM-Newton observations of the NS-picture.html galactic supernova remnant CTB 109 (G109.1-1.0), *ApJ* **617**, 322.
- Seaquist, E. & W. Gilmore (1982) Radio and X-ray observations of compact sources in or near supernova remnants, *ApJ* **378**, 378.
- Seward, F. D. & Z.-R. Wang (1988) Pulsars, X-ray synchrotron nebulae, and guest stars, *ApJ* **332**, 199.
- Seward, F. D., F. R. Harnden Jr. & D. J. Helfand (1984) Discovery of a 50 millisecond pulsar in the Large Magellanic Cloud, *ApJ* **287**, L19.
- Seward, F., P. Charles & A. Smale (1986) A six-second periodic X-ray source in Carina, *ApJ* **305**, 814.
- Seward, F., P. Gorenstein & R. Smith (2006) Chandra observations of the X-ray halo around the Crab Nebula, *ApJ* **636**, 873.
- Seward, F. D., et al. (2000) ROSAT observations of the Vela Pulsar, *ApJ* **536**, 948.
- Seward, F., et al. (2003) A compact central object in the supernova remnant Kesteven 79, *ApJ* **584**, 414.
- Slane, P., et al. (2004) New constraints on the structure and evolution of the pulsar wind nebula 3C 58, *ApJ* **616**, 403.
- Tatematsu, K., et al. (1987) Molecular Clouds in the vicinity of the semicircular supernova remnant G109.1-1.0, *A&A* **184**, 279.
- Walter, F. M., S. J. Wolk & R. Neuhauser (1996) Discovery of a nearby isolated neutron star, *Nature* **379**, 233.
- Weisskopf, M., et al. (2000) Discovery of spatial and spectral structure in the X-ray emission from the Crab Nebula, *ApJ Lett* **536**, L81.
- Williams R. M., et al. (2005) Supernova remnants in the Magellanic Clouds. V. The complex interior structure of the N206 supernova remnant, *ApJ* **628**, 704.
- Kennel, C. F. & F. V. Coroniti (1984) Magnetohydrodynamic model of Crab nebula radiation, *ApJ* **283**, 710.
- Lattimer, J. M. & M. Prakash (2004) The physics of neutron stars, *Science* **304**, 536.
- Oppenheimer, J. R. & G. M. Volkoff (1939) On massive neutron cores, *Phys Rev* **55**, 374.
- Pacini, F. & M. Salvati (1973) Evolution of the Crab Nebula, *ApL* **13**, 103.
- Page, D. (2009) <http://www.astroscu.unam.mx/neutrones/NS-picture.html>.
- Rees, M. J. & J. E. Gunn (1974) The origin of the magnetic field and relativistic particles in the Crab Nebula, *MNRAS* **167**, 1.
- Thompson, C. & R. C. Duncan (1995) The soft gamma repeaters as very strongly magnetized neutron stars – I. Radiative mechanism for outbursts, *MNRAS* **275**, 255.
- Van Riper, K. A., R. I. Epstein & G. S. Miller (1991) Soft X-ray pulses from neutron star glitches, *ApJ* **381**, L47.
- Yakovlev, D. G. & C. J. Pethick (2004) Neutron star cooling, *Ann Rev Astron Ap* **42**, 169.

Reviews, catalogues, popular

- Gaensler, B. M. & P. O. Slane (2006) The evolution and structure of pulsar wind nebulae, *Ann Revs Astron Ap* **44**, 17.
- Hobbs, G. B. & R. N. Manchester (2006) ATNF pulsar catalog, <http://www.atnf.csiro.au/research/pulsar/psrcat/>.
- Kaspi, V., M. Roberts & A. Harding (2006) Isolated neutron stars, in *Compact Stellar X-ray Sources*, 249, edited by W. Lewin & M. Van Der Klis, New York: Cambridge University Press.
- Michanowsky, G. (1977) *The Once and Future Star*, New York: Hawthorn Books.
- Seward, F. D. (1978) A trip to the Crab Nebula, *J Br Interplanet Soc* **31**, 83.
- Seward, F. D., et al. (2006) Chandra supernova remnant catalog, <http://hea-www.harvard.edu/ChandraSNR/>.

Theory

- Gold, T. (1968) Rotating neutron stars as the origin of the pulsating radio sources, *Nature* **218**, 731.

Chapter 10

Cataclysmic variable stars

10.1 Introduction

Cataclysmic variables are remarkably similar to the low-mass X-ray binaries (LMXBs) described in Chapter 11 but are significantly less luminous in X-rays, particularly at the higher X-ray energies of the first X-ray satellites. Indeed, only one previously known CV, EX Hya, was found in the Uhuru survey of the X-ray sky. But these objects have a long history that goes back well before the era of X-ray astronomy, as they include among their number both dwarf novae and novae. As we shall see in this chapter, *dwarf novae* and *novae* are powered by fundamentally different processes than occur in *supernova* events. Supernovae represent the final, and irreversible, moments in the lives of massive stars, when they collapse rapidly under gravity and then explode catastrophically. For any given object it happens once, whereas nova and dwarf nova eruptions can and do recur. Indeed, it is hypothesised that all novae recur, but the typical recurrence time is long: at least hundreds, perhaps thousands, of years.

Cataclysmic variables are *interacting binaries* similar to LMXBs, except that the compact object is a white dwarf, accreting material from its (usually) cool, late-type companion star in a short period (approximately hours) binary system. They are one of the few classes of object in this book that were known and observed prior to the twentieth century. Novae have been known to mankind throughout history, and dwarf novae were first recorded in the mid-nineteenth century. For a few

weeks the nearest and brightest novae can become important naked-eye objects. Indeed, Nova Cygni in 1975 reached second magnitude at its peak and, for a time, dramatically changed the appearance of the constellation of Cygnus, the Swan, for a few weeks in summer 1975 (an event which the authors remember well!). Novae are now known to be the result of unstable thermonuclear burning on the surface of a white dwarf (of material provided by its companion star), and accretion energy contributes little to the outburst. The mass transfer onto the white dwarf occurs gradually over a long period of time between nova outbursts, which take place when about $10^{-4} M_{\odot}$ of material has been accumulated on the surface. The low level of mass transfer does release accretion energy, but X-ray emission from this process is weak.

It was the irregular but continuing eruptions displayed by dwarf novae (such as U Gem and SS Cyg) that led to their *cataclysmic* designation, a term now widely used for all accreting white dwarfs in interacting binary systems (see Fig. 10.1a for a typical dwarf nova light curve). However, their physical nature was not understood until the pioneering spectroscopic studies of Bob Kraft in the late 1950s which revealed their binary signature. Because of the accessibility of a significant number of CVs to amateurs equipped with small telescopes, they are popular targets for variable star observers. Indeed, this is one field of astronomy where, even today, amateur groups play an important role in supporting and stimulating professional observations (see <http://cbastro.org/>). Although the dwarf nova outburst is much less

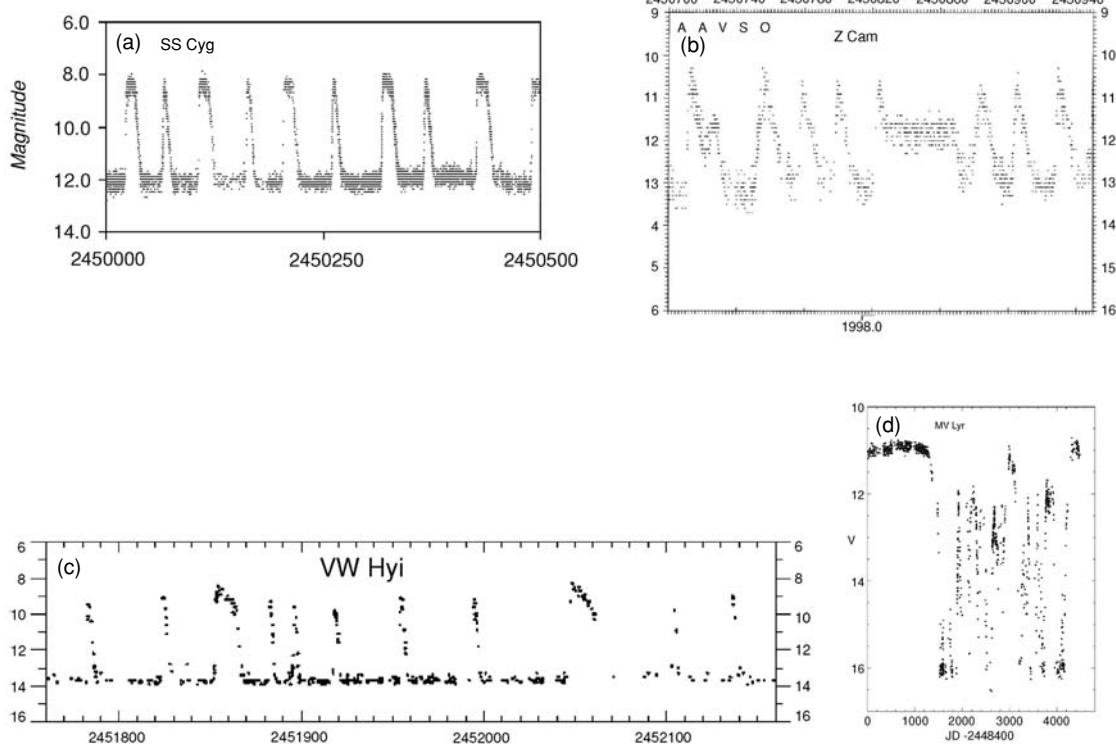


Fig 10.1 Typical light curves of cataclysmic variables. Since the dominant source of light in a CV is the accretion disc, this demonstrates the great variety of behaviour of which such a disc is capable. (a) Classical dwarf nova outbursts of SS Cyg during 1995–1996. (b) How a disc can become stuck halfway between quiescence and outburst (this shows Z Cam entering a standstill for about 2 months). (c) The SU UMa system, VV Hydri, displaying a mix of normal and occasional superoutbursts. (d) Anti-dwarf nova MV Lyr which is normally bright but occasionally drops to a low state (adapted from AAVSO historical light curves at <http://www.aavso.org/vstar/historical.shtml>).

dramatic than a typical nova explosion (novae can and do reach naked-eye magnitudes, whereas no dwarf nova has ever been recorded by the unaided eye), they occur more frequently (some objects producing outbursts every week or two) and last for much shorter periods of time (usually less than 10 days or so). Table 10.1 summarises the observed properties of typical nova and dwarf nova outbursts.

There is a great variety in the nature of dwarf nova outbursts, as indicated by the sample shown in Fig. 10.1.

A classical dwarf nova, such as U Gem or SS Cyg, increases in brightness by 3–5 magnitudes within a day and then decays on a timescale of about 10 days. There are typical times between outbursts for any particular object, but these do vary erratically, which makes it very difficult to study

them through conventionally assigned telescope time.

The figure also indicates the behaviour of other types of dwarf novae. The short period SU UMa systems appear at first sight to be like other classical dwarf novae in their outburst patterns. But much more rarely than their normal outbursts, they deliver an exceptionally long outburst which has become known as a *superoutburst*.

The Z Cam stars also appear to erupt normally like other dwarf novae, but occasionally they get stuck in *standstills*, sometimes for months on end, at a brightness level below their normal outburst maximum but well above normal quiescence. Finally, there are the *anti-dwarf novae*, or VY Scl stars, which spend most of their time in an outburst state and occasionally dip into quiescence for a few days.

Table 10.1 Observed properties of cataclysmic variables.

Type	$\log L_X$ (erg s $^{-1}$) (0.2–10 keV)	kT (keV)	Donor star	Examples	P_{orb} (hours)	Notes
Dwarf Nova	29–32	0.025, few-10 (multi-T)	G-M	U Gem SS Cyg	4.2 6.6	$\sim 100 \times$ range in \dot{M}, L quiescent to outburst
Novalike	31–32	0.25–0.5		UX UMa IX Vel	4.7	Link to SSS (e.g. VY Scl)
Nova	33–38	0.2–15 (multi-T)		DQ Her V1500 Cyg	4.6 3.4	Change in R_{phot} Hard shock with ISM
‘Old’ Nova	30–33	>1		V603 Aql GK Per	47.9	Similar to quiescent DNe
AM CVn	28–30.5	<0.25	Degenerate	AM CVn V407 Vul		All have $P < 60$ min
Symbiotic	30–33	0.25–1.3	Red giant			Link to SSS Hard X-rays from colliding winds
Polar	32–33	$\sim 0.03, \sim 5$ (multi-T)	$\sim M$	AM Her EF Eri	3.1 1.35	Sync. rotation Most are below period gap
Intermed. polar	32–33	few-30		TV Col EX Hya AE Aqr	5.5 1.6 9.9	Some also have soft component
Supersoft source	36–38	0.01–0.08	Subgiant	CAL83 RXJ0513-69		$P \sim 2$ hours to 4 days Inverted mass ratio

Note. For the most detailed technical review of CV properties, see Warner (1995).

There are also objects known as *novalikes* because they look like a nova would long after its eruption but in this case they have not actually been observed in a nova state. Nor do these objects display typical dwarf nova outbursts. The classical example of this is UX UMa.

For a long time this wide variety of behaviour was considered to be an impenetrable forest of different types which may or may not relate to one another. However, observations at X-ray and UV wavelengths have enabled major inroads to be made into the puzzles of these low-mass interacting binaries (key discoveries are listed

chronologically in Table 10.2). CVs can basically be divided into magnetic and non-magnetic systems, according to the properties of the white dwarf. In non-magnetic CVs, the accretion disc dominates the emitted radiation, whereas if the magnetic field is strong enough, the flow of matter onto the white dwarf is dominated entirely by magnetic effects.

Why are CVs important objects in astronomy? Firstly, the visible light we see is normally dominated by the accretion process itself, in particular, the *accretion disc* that forms around the white dwarf. This disc is heated by its internal viscosity

Table 10.2 | Chronology of cataclysmic variables from X-ray/UV observations.

Year	Discovery
1972	Hard X-rays detected from EX Hya (Warner, 1972)
1974	Soft X-rays detected from SS Cyg in outburst (Rappaport <i>et al.</i> , 1974)
1977	Large, variable polarisation discovered in soft X-ray source AM Her (Tapia, 1977)
1977	Term <i>polars</i> coined by Krzeminski and Serkowski (1977) for AM Her systems
1978	Soft X-rays detected from SS Cyg in outburst and quiescence (Heise <i>et al.</i> , 1978)
1980	14.3-min X-ray pulsations discovered in H2252-035 (Patterson & Price, 1981)
1981	<i>Supersoft sources</i> (SSS) found by Einstein in LMC (Long <i>et al.</i> , 1981)
1981	Term <i>intermediate polars</i> coined by Krzeminski (1981) for asynchronous WD rotation
1983	Einstein X-ray survey of CVs (Córdova & Mason, 1984)
1991	IUE spectra show that U Gem WD cools between outbursts (Kiplinger and colleagues)
1992	Steady (super-Eddington) nuclear burning WD model for SSS (van den Heuvel <i>et al.</i> , 1992)
1996	AE Aqr ‘magnetic propellor’ model (Eracleous & Horne, 1996)
2002	RXJ0806+15 proposed as shortest known P_{orb} , 5.4 mins (Ramsay <i>et al.</i> , 2002)

to typically 5000–10 000 K, whereas its cool, mass-losing companion is less massive than the Sun and hence faint. The white dwarf may be hot, but it is very small (comparable to the Earth in extent) and contributes little except in the UV.

Secondly, they are numerous in our Galaxy. The majority of the several hundred CVs that have been catalogued are within a few hundred parsecs, making them easy to study. In particular, their proximity means they suffer little interstellar extinction, and hence their spectra can be studied over a wide range of wavelengths, including soft X-rays and UV. Here the inner part of the accretion disc and the white dwarf dominate, at temperatures exceeding 20 000 K.

CVs therefore allow the most detailed observations of accretion physics at work, processes that are now recognised as being significant on virtually all scales, from star and planetary formation (the proto-star accretes material from its surrounding molecular cloud) to accretion onto supermassive black holes in the centres of AGN (see Chapter 14). Unfortunately, in most of these areas, the geometry and mass accretion processes are very poorly constrained, there are a variety of emitting sources that are complex to unravel and the relevant timescales can be uncomfortably long compared to human lifetimes. However, CVs are interacting binaries with a geometry that is well defined by orbital mechanics. This provides an excellent astronomical

‘laboratory’ in which to study the physics of accretion onto compact objects, in particular the nature of the viscosity which acts within the disc. CVs thereby provide constraints on the basic physical properties of accretion discs, which can then be applied in many other fields. At the same time, more can be learned about the end points of stellar evolution which all stars must eventually go through.

10.2 Geometry of accretion in CVs

The basic geometry of CVs, shown schematically in Fig. 10.2, was recognised in the 1960s. The work of Bob Kraft and co-workers in the late 1950s demonstrated that dwarf novae contained accretion discs by exploiting the benefits afforded by observing eclipsing binaries. In such objects we are fortunate to be located in the orbital plane of the binary. Time-resolved optical spectra demonstrate the presence of an *extended object*, namely, the accretion disc (see Fig. 10.3).

The clearly double-peaked structure arises because of the motion of material in and from the disc itself, as shown in the figure. This interpretation was verified by the time-resolved spectra obtained through eclipse, where the companion

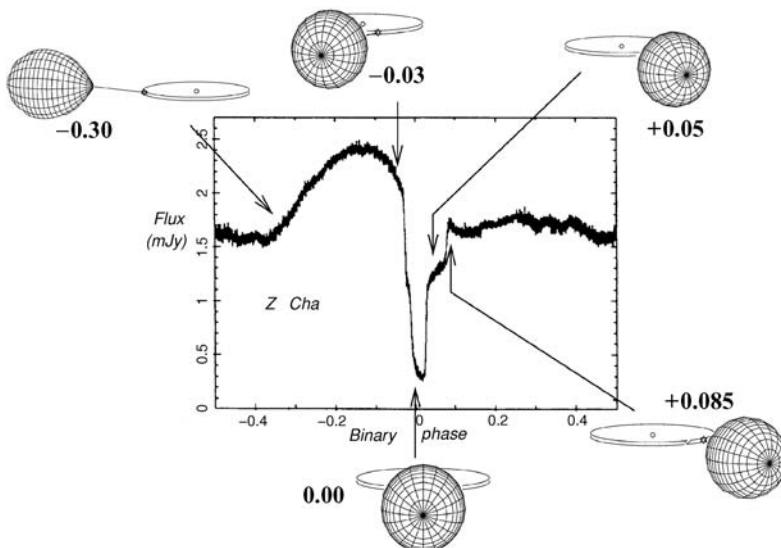


Fig 10.2 Basic geometry of a CV, in which a low-mass normal star transfers material into an accretion disc surrounding the white dwarf. The mass-transfer stream produces a bright spot at the edge of the disc, which is seen as the large hump in the light curve. The binary and its resulting light curve are shown at different orbital phases using quiescent (low-state) observations of the eclipsing dwarf nova Z Cha as an example (adapted from Marsh, 2001).

star occults first one side of the disc then the other. We can actually see the accretion disc.

10.2.1 Accretion disc spectrum

The main component observed in a typical dwarf nova, whether in quiescence or outburst, is the accretion disc itself, as shown in the representation in Colour Plate 28. Matter is gradually transported through the disc onto the white dwarf, in the process being heated to higher temperatures and making the inner disc a major source of UV emission. Furthermore, only about half of the

total accretion luminosity is deposited in the disc; the other half is released in the region between the disc and the white dwarf, where the accreting matter is finally brought to a halt. This is known as the *boundary layer*, is even hotter and radiates predominantly in the EUV and soft X-ray regions.

The disc is therefore hotter in its inner regions, and its temperature decreases towards the outside. The picture of the disc obtained depends on which wavelength region or colour it is observed in, as described in more detail in Box ‘Multi-colour disc spectrum’. The extended, cooler outer regions

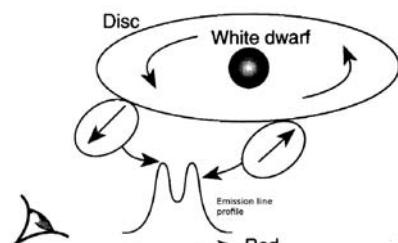
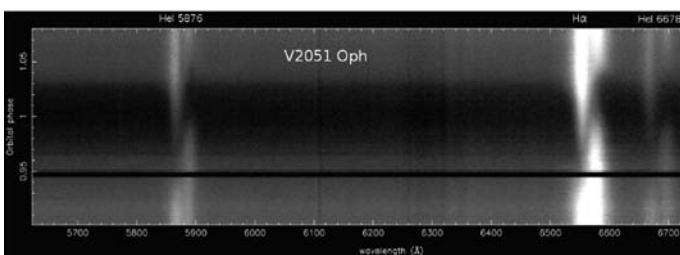


Fig 10.3 Time-resolved optical spectra obtained with the Keck Telescope through an eclipse in V2051 Oph. Time goes from bottom to top (plotted as binary phase), and the eclipse (centred at phase 1.0) is visible as the broad horizontal band (the short, sharp band near phase 0.95 is a data gap). This red section of the spectrum is dominated by the bright H α emission line at 6563 Å. The H α profile shows the classical double peaks of an accretion disc viewed at high inclination (the plane of the disc is close to our line of sight). The schematic shows how this arises from the rotation of the disc about the white dwarf. The projection on the sky leads to the dominant emission coming from each side of the white dwarf, one of which is approaching us (blue shifted), whilst the other is receding (red shifted), thereby giving rise to the two peaks. The time resolution in this ‘trailed’ spectrum through eclipse confirms this geometry since the (essentially dark) companion star obscures first the blue peak (starting phase 0.95) then the red peak (starting phase 1.0), and then uncovers them in the same order (see Fig. 10.2) (spectra from Danny Steeghs).

of the disc (and the companion star itself) are usually brighter in the IR, whereas the inner disc dominates in the UV. And because the inner disc is very hot, the overall colour of the system is very blue. It is this characteristic above all others that is used most frequently to distinguish CVs from other objects.

While this basic model of a CV is indeed very similar to that of an LMXB, the large (factor $\sim 10^3$) difference in radius of the white dwarf compared to a neutron star means that for a given mass transfer rate, its luminosity will be that much smaller. Consequently, non-imaging X-ray satellites of the 1970s (Uhuru, Ariel 5 and several others) were used to search for the hard (>1 keV) X-rays that had been expected from the white dwarf surface. Only EX Hya and AM Her (both magnetic systems; see later sections) are in the Uhuru X-ray catalogues, but even the latter required the discovery of strong soft X-ray emission and a more precise location on the sky for it to be firmly identified. The field of X-ray-emitting CVs did not open up until the greatly enhanced sensitivity of the first imaging X-ray satellite, the Einstein Observatory, which found CVs in large numbers and with $L_X \sim 10^{30-33}$ erg s $^{-1}$ (Córdova & Mason, 1984).

10.2.2 The companion star

This analysis of accretion disc spectra is well supported by observations of CVs where the

continuum emission almost always exhibits a hot, blue, non-stellar component. With orbital periods of hours and a white-dwarf primary star (of mass $\leq 1 M_\odot$), it is easy to show that the companion star must be of low mass. (There just is not room enough in the system for a massive star.) Furthermore, this cool, low-mass star (usually referred to as the *secondary star* or *companion*) must be filling its Roche (equipotential) surface and hence transferring material into the hot accretion disc through its inner Lagrangian point (see Fig. 10.2) that gives rise to the emission lines and blue continuum. (For details [and equations] of the Roche geometry in interacting binaries, see Warner, 1995, the primary reference source for CV properties).

The source of matter for the disc, which must be continuously replenished, is the companion star. It is usually not too different in size or appearance from that of a main sequence star of comparable mass, except for the fact that it is transferring material to the compact object that it is orbiting. As this matter falls in the potential well of the white dwarf, it heats the disc, which subsequently radiates.

10.3 | Dwarf nova outbursts

A typical CV outburst (Fig. 10.1) involves it brightening by as much as a factor 100 within a day,

Box 10.1 | Multi-colour disc spectrum

The luminosity and temperatures produced in CVs can be estimated as follows. A white dwarf accreting matter at a rate \dot{M} will produce a luminosity $L = G M_{WD} \dot{M} / R_{WD}$ through the release of gravitational potential energy. For a white dwarf of mass $M_{WD} = 1 M_\odot$ and radius $R_{WD} = 10^4$ km accreting at $\sim 10^{-10} M_\odot \text{yr}^{-1}$ (a typical value from calculations of binary evolution), the total luminosity is $\sim 10^{33}$ erg s $^{-1}$. Assuming that this energy is released as blackbody radiation ($L = 4\pi R^2 \sigma T^4$) from the white dwarf surface, then $T \sim 30\,000$ K, and so CVs should be copious sources of UV radiation. Indeed, most CVs do show very blue colours rising into the UV, implying that the bulk

of the emission will lie at shorter wavelengths. Assuming further that the accreting material free-falls onto the white dwarf surface, then it should reach impact velocities ~ 5000 km s $^{-1}$, producing temperatures $\sim 10^9$ K, i.e. very hard X-rays.

To calculate the emitted spectrum of an accretion disc, we must look more closely at its properties. A disc must be formed because the transferred matter from the donor carries orbital angular momentum and hence cannot (in the absence of other forces) accrete directly onto the white dwarf. It therefore enters an orbit around the white dwarf. The disc's internal viscosity results in the transport of angular momentum outwards

and material inwards. A physical explanation of this viscosity remains one of the major unsolved problems in modern astrophysics and is a very active research area today. Nevertheless, our lack of understanding of viscosity is concentrated into the parameter $\alpha = \nu/c_s H$ (ν is the effective turbulent viscosity, c_s the sound speed and H the half-thickness of the disc), which was used in a seminal paper by Shakura and Sunyaev (1973) to produce so-called α discs, which have been the starting point for all accretion disc studies ever since. In the steady state, the disc spectrum does not depend on α , but as might be expected, the response time of the disc to changes in \dot{M} does. This means that it is through studying the outburst properties (e.g. durations and recurrence times) of dwarf novae that the physics of disc viscosity can be constrained (see e.g. the review by Lasota, 2001).

Pringle (1981) summarised the Shakura-Sunyaev approach, from which the disc spectrum can be calculated. A circular, thin disc is divided

radially into annuli of radius r , each of which is assumed to radiate as a blackbody at temperature $T(r)$, where

$$T(r)^4 = T_{\max}^4 (r/R_{WD})^{-3} [1 - (R_{WD}/r)^{1/2}] \quad (10.1)$$

and T_{\max} is the disc's maximum temperature, given by

$$T_{\max} = 41000 \left(\dot{M} M_{WD} / R_{WD}^3 \right)^{0.25} K, \quad (10.2)$$

where \dot{M} is in units of $10^{16} g s^{-1}$, M_{WD} is in solar masses and R_{WD} is in units of 10^4 km. The total spectrum emitted by the disc is then the sum of the contributions from each of these annuli, as demonstrated in Fig. 10.4, the result of which is a continuum with a very blue colour. Such a spectrum has the characteristic UV excess shown by the majority of CVs. However, one must take into account the much cooler temperatures of the outer-disc regions (and the companion star), so it is necessary to observe from the near-IR into the far-UV to fully characterise the CV spectrum.

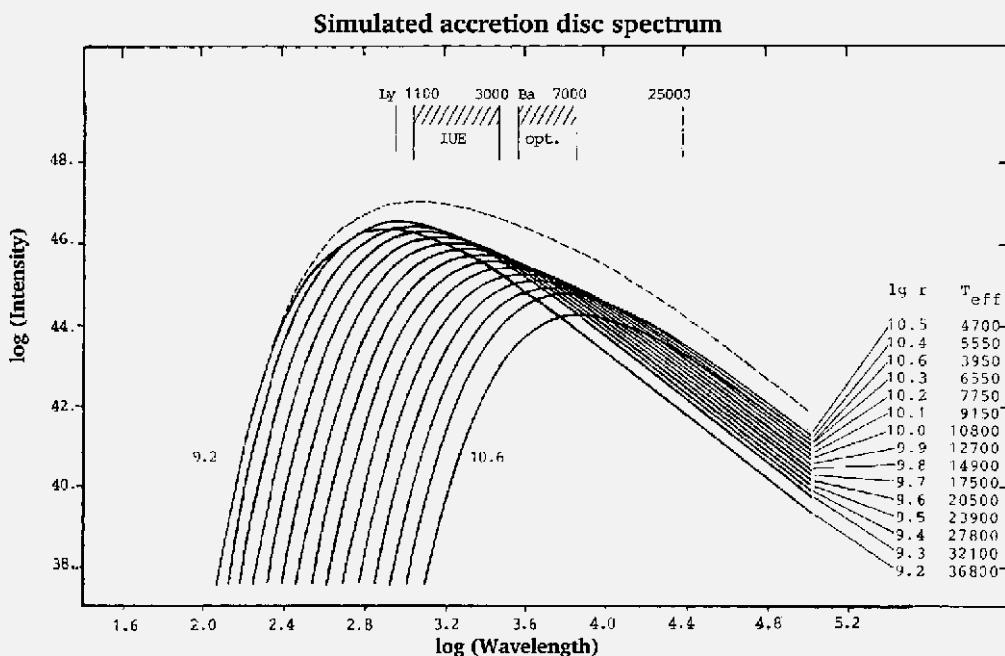


Fig 10.4 The dashed line shows a theoretical spectrum of an accretion disc computed by adding the contributions of individual rings or annuli in the disc, each of which is assumed to radiate as a blackbody. The contribution of each ring is labelled with its radius and temperature (coolest on the outer edge) (adapted from La Dous, 1989).

Box 10.2 | Disc thermal instability

In the great majority of CVs, the accretion discs consist mainly of hydrogen, and it is the physical properties of hydrogen that account for the instability used in the DIM. First worked out by Hoshi, Pringle, Meyer and Meyer-Hofmeister, the ‘limit cycle’ is shown schematically in Fig. 10.5. Beginning in the quiescent state, matter is transferred into the disc from the companion at a slow but essentially steady rate. The disc’s surface density therefore gradually increases, along with the disc temperature. This continues until it is hot enough for hydrogen to begin to be ionised (around 6000 K, point ‘A’ in the diagram), which leads to a dramatic change in behaviour. This is because ionised hydrogen increases the disc’s opacity and viscosity, moving it rapidly to the hot, viscous state (upper arm in the ‘S curve’).

In this hot outburst state, the higher viscosity leads to a much greater inward flow of matter onto the white dwarf, which exceeds the rate at which it is replenished by the companion star. Hence the disc loses density and moves downwards towards point ‘B’, where the hydrogen begins to recombine, leading it back to the cool, quiescent state.

remaining at that level for a day or two and then returning gradually to its previous level over a period of about a week. The only component that can possibly vary in this way is the disc. There is no known mechanism by which the (essentially normal) companion star can brighten significantly, and the white dwarf presents too small an area to radiate this much luminosity. Since the early 1970s and well into the 1980s, there were two competing models to account for the origin of the dramatic change in brightness of the accretion disc, which led to entertaining and lively debates amongst the protagonists.

The first model links the outburst to the rate at which mass is being transferred from the companion star. If this were to increase for some reason, then there would be a corresponding increase in the luminosity of the disc. This is called the *mass-transfer model*.

The other explanation for CV outbursts assumes that the mass transfer rate from the

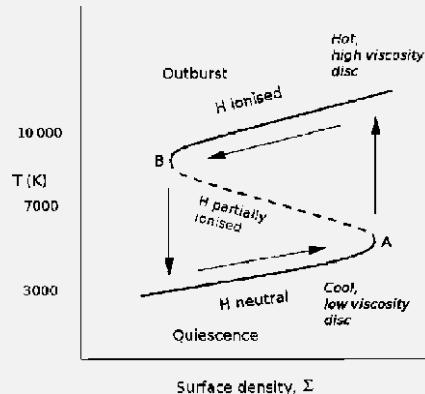


Fig 10.5 A dwarf nova disc moves between states of quiescence, when the disc is in a cool, low-viscosity state, and outburst, when it is hot and viscous, following an S curve. The addition of material from the companion heats up the disc until it reaches point ‘A’, when it is hot enough for hydrogen to begin to ionise. This causes the viscosity and temperature to increase dramatically, and the disc moves up to the hot, high-viscosity state. The disc now cools until point ‘B’, when hydrogen begins to recombine and the disc transitions back to the cool, quiescent state (adapted from an original by S. Mineshige).

secondary is roughly constant with time. The disc accumulates more and more material whilst changing little in appearance, until, suddenly, the internal structure of the disc changes (its viscosity must increase), and it brightens dramatically. This is the *disc-instability model* (DIM).

Much observational and theoretical effort went into attempting to distinguish between these two models, and it is now generally accepted that even though the mass-transfer rate from the donor can vary substantially, the DIM provides by far the best explanation of the many outburst observations. In particular (as is clear in Colour Plate 28), the bright spot on the edge of the disc (which is easily visible in dwarf nova light curves) would be expected to vary with the mass-transfer rate changes. This is not seen, and instead, the luminosity release in outburst is compatible with the DIM. Furthermore, the DIM explains many of the properties of dwarf nova quiescence to outburst cycles (see Box 10.2).

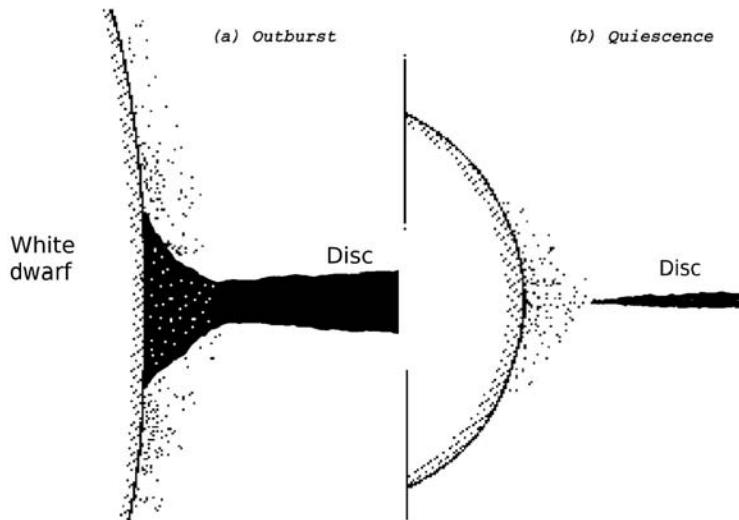


Fig 10.6 Schematic view of the inner disc–white dwarf regions of a dwarf nova in (a) outburst and (b) quiescence. In quiescence, matter accretes slowly onto the white dwarf, but it is hot and optically thin, hence producing mostly hard X-rays. In outburst the much higher accretion rate through the disc (due to its higher viscosity) leads to an optically thick, soft X-ray-emitting region. This can still, however, be surrounded by hotter, optically thin regions (adapted from Patterson & Raymond, 1985).

10.4 | X-rays from dwarf novae

It was the HEAO-1 discovery in the late 1970s of very soft X-rays (<0.2 keV) from SS Cyg and U Gem during outburst that led to major advances in CV studies. Such soft X-rays implied a low-temperature (10–50 eV, or equivalently $1\text{--}5 \times 10^5$ K) blackbody. This soft component was associated with the disc–white dwarf boundary layer region (Fig. 10.6). In quiescence, however, much lower accretion rates produce a hard X-ray spectrum because most of the energy is released close to the white dwarf surface in a hot, optically thin plasma. There may also be an extended, very hot ($\sim 10^8$ K) corona because the cooling at these temperatures is very inefficient, and so the plasma expands.

10.4.1 Testing the disc instability model

A key prediction of the disc instability model just described was that the outburst would begin when the disc's outer regions made the transition to the hot, high-viscosity state (Fig. 10.5). This transition would then move inwards through the disc as a ‘heating wave’, eventually reaching the inner disc and at that point increasing the mass transfer onto the white dwarf. This would result in the dwarf nova outburst beginning first in the optical (where the outer disc dominates; see Fig. 10.4) and then being followed by the UV (from the inner

disc and white dwarf surface) – the so-called UV delay. This prediction was confirmed via simultaneous ground-based and International Ultraviolet Explorer (IUE) observing campaigns in the 1980s, which measured an ~ 0.5 -day delay of the UV peak relative to the optical. (It is worth noting that this is an area where amateur observers can play a major role because of the inherent unpredictability of dwarf nova outbursts. Amateurs provide the continuous monitoring of target CVs, which is crucial in catching objects such as dwarf novae in particular states. Of course, the spacecraft, too, must be able to respond rapidly to such requests, and observations like these now come under the special category designated TOO observations.)

But how does the X-ray emission (arising from the accretion onto the white dwarf) fit in with the UV delay? From the basic physical model of the inner disc regions during outburst and quiescence (Fig. 10.6), it would be expected that early in the outburst (which starts in the optical in the outer disc and before the accretion rate onto the white dwarf has had a chance to change), the hard X-rays should initially be unaffected. But as the outburst progresses and the accretion rate onto the white dwarf increases, then the hard X-ray output should decrease and be replaced by soft X-rays from the optically thick boundary layer.

Wide multi-wavelength coverage (optical, EUV and hard X-rays) of a dwarf nova outburst was

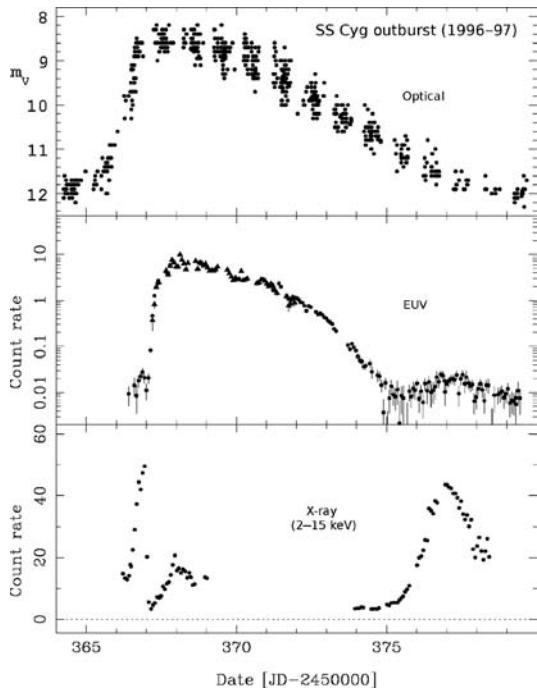


Fig 10.7 Outburst of the dwarf nova SS Cyg from the mid-1990s which was observed simultaneously at optical, EUV and X-ray wavelengths. This reveals the complex behaviour of the soft and hard X-ray components as the outburst progresses (adapted from Wheatley et al., 2000).

finally achieved (Fig. 10.7) in the 1996–1997 outburst of SS Cyg (Wheatley et al., 2000), giving results that were clearly much more complex than anticipated.¹ The UV delay is clearly visible, but the hard X-rays initially increased before declining to be replaced by the EUV emission. What is believed to be happening here is that at the start of the (optical) outburst (which begins at the outer edges of the disc), the inner regions are truncated (i.e. they contain very little material). These inner regions are optically thin and initially emit hard X-rays while matter fills up the inner disc down to the white dwarf surface. Once this boundary layer becomes optically thick, the hard X-rays decline and are replaced by strong EUV emission. (For a more detailed discussion, see Lasota, 2001.) UV observations also suggest that the white dwarf

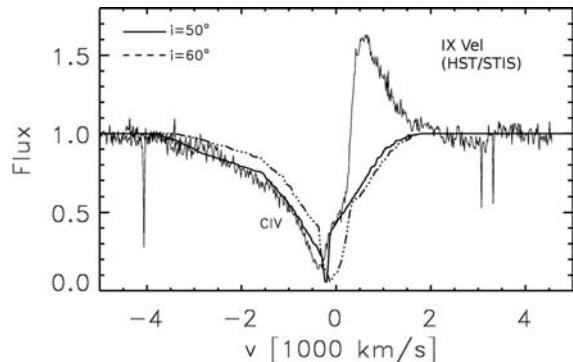


Fig 10.8 High-resolution HST spectrum of IX Vel showing the P Cyg profile exhibited by CIVλ 1550 compared with theoretical profiles (for two different inclination angles) of the absorption component only. The outflow velocity inferred in the disc wind is 3500 km s^{-1} (adapted from Proga, 2003).

cools after each outburst, an effect that is likely confined to an equatorial band where the interface with the disc is located and most of the additional accretion energy is deposited.

10.4.2 Accretion disc winds

The inner disc and white dwarf surface have temperatures in the range $\sim 1\text{--}5 \times 10^4 \text{ K}$, so it is UV observations that have produced an understanding of their properties since this is where the peak of the emission lies. This was demonstrated by early IUE spectra which revealed asymmetric and even P Cygni profiles of the emission lines. More recent HST high-resolution spectra (Fig. 10.8) of the brightest novalike CV, IX Vel, provide strong evidence for the existence of fast streams of outflowing material, now described as an *accretion disc wind*. The velocities inferred for this outflowing gas (approximately a few thousand kilometers per second; see Fig. 10.9) are comparable to the white dwarf escape velocity. Such winds are believed to be significant in discs on a wide variety of scales.

10.4.3 Rapid variability

10.4.3.1 Optical and X-ray flickering

Chaotic optical and X-ray variability on timescales of minutes (and usually referred to as *flickering*) is

¹ AAVSO amateurs phoned Pete Wheatley to inform him that the SS Cyg outburst had begun, but to trigger his RXTE TOO, he had to cycle to his office to send an e-mail. However, he was so excited at the news that he had an accident. Nevertheless he continued on to send the e-mail and successfully trigger the TOO before attending to his wounds!

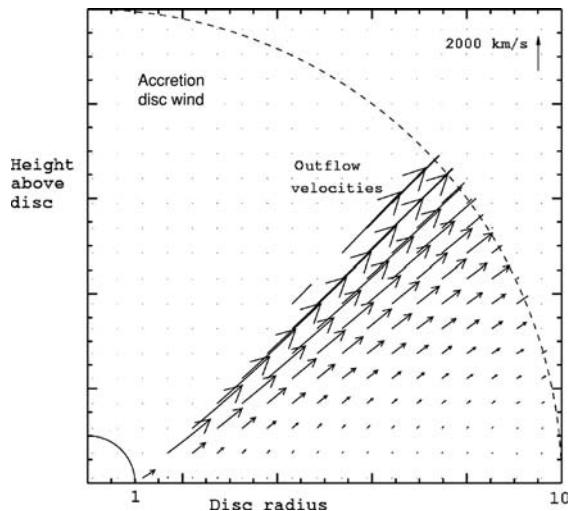


Fig 10.9 Irradiation of the disc by accretion energy released close to the white dwarf drives an outflow of material, a ‘disc wind’ (adapted from Proga *et al.*, 1999).

seen in most CVs. Eclipsing systems (like U Gem) provide an opportunity to locate the source of this variability, and it is observed that when the bright spot is eclipsed, the level of flickering reduces. This means that some component of flickering must arise in the bright spot (where the accretion stream impacts the disc; see Fig. 10.2). However, polars (see Section 10.6) have no disc at all (and hence no bright spot), yet they also flicker, probably due to shocks in the accretion column onto the white dwarf. Also, in novalike CVs (where the disc stays permanently in the hot, outburst state), the optical flickering correlates with X-ray variability, which means that it must arise in the *inner* disc (as the outer disc is too cool to produce X-rays). Apparently, flickering originates throughout the accretion process.

10.4.3.2 Quasi-periodic oscillations

All CVs display apparently regular pulsations with periods of tens of seconds (summarised in Warner’s 1995 review), which are quite small (<1% amplitude) in the optical but larger in X-rays. They are not strictly periodic (hence the name *Quasi-periodic oscillations*, or QPOs), but some are stable over long timescales. Remarkably, QPOs can return at essentially the same value after many years; e.g. Warner and Nather (1972) first observed

29-s QPOs in UX UMa, and then a quarter of a century later, Knigge *et al.* (1998) detected UV/optical QPOs at almost the same period with HST.

Nevertheless, these pulsations are not coherent (as for the rigidly rotating components in magnetic CVs), which is why they are called QPOs. Most optical QPOs are much less stable in both period and phase, and the X-ray QPOs seen in SS Cyg (~ 10 s) and U Gem (~ 25 s) at outburst are of much larger amplitude than those in the optical. This infers that the primary QPO is in the X-ray band, and the optical QPO is due to reprocessing of the X-rays in the inner disc region. This suggests that the physical mechanism driving QPOs in CVs may be similar to the fast X-ray QPOs occurring in the much more luminous LMXBs. These are believed to arise where the inner accretion disc interfaces with the magnetosphere of the rapidly spinning compact object.

10.5 How do CVs form and evolve?

The general idea of how low-mass interacting binaries (such as CVs) were formed was put forward by the late Bohdan Paczynski in the mid-1970s (see Chapter 11). Double stars are created in large star-forming regions (e.g. the Orion Nebula) with the gravitational collapse of clouds of gas and dust. Such double stars have wide separation (several AU) and orbital periods of years, very different from what we see in CVs. This is because of the problem of disposing of the angular momentum inherited from the initial gas cloud out of which the stars form. However, it is normal stellar evolution that will radically change the picture.

Both stars begin their lives on the main sequence (like the Sun), but the more massive component will evolve (burn its fuel) more rapidly and start to expand into a red giant (Fig. 10.10). The expansion is so great that the atmosphere of the more massive star actually swells to embrace the other star, which then spirals in towards the core of the red giant as a result of atmospheric drag (with the atmosphere carrying off the angular momentum of the binary). You can almost imagine the core of the red giant (which itself is

Interacting binary evolution

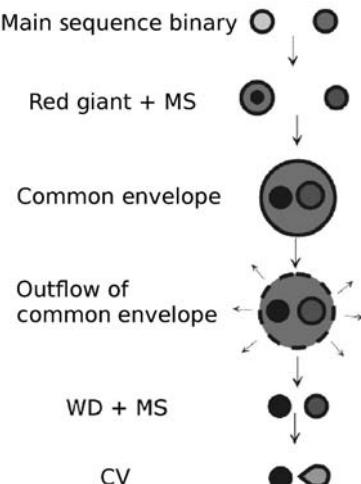


Fig 10.10 (top) Formation of a CV from a pair of main sequence stars. Normal stellar evolution of the larger of the pair into a red giant leads to a common envelope phase and the spiral-in of the remaining normal star towards the white dwarf core of the evolved star. This process will release energy that will drive off the envelope, leaving a white dwarf in a short-period orbit about its stellar companion. Subsequent evolution of this star will eventually lead to mass transfer onto the white dwarf and the beginning of its CV phase.

collapsing towards the white dwarf stage) and the companion star orbiting each other actually *inside* the overall envelope of the red giant. Note though that this envelope is very tenuous and dissipates

rapidly as a result of the spiral-in, leaving a short-period binary with the companion star orbiting the white dwarf. Eventually the companion star will evolve to fill its own Roche lobe, and mass transfer onto the white dwarf will then turn on. A CV is born.

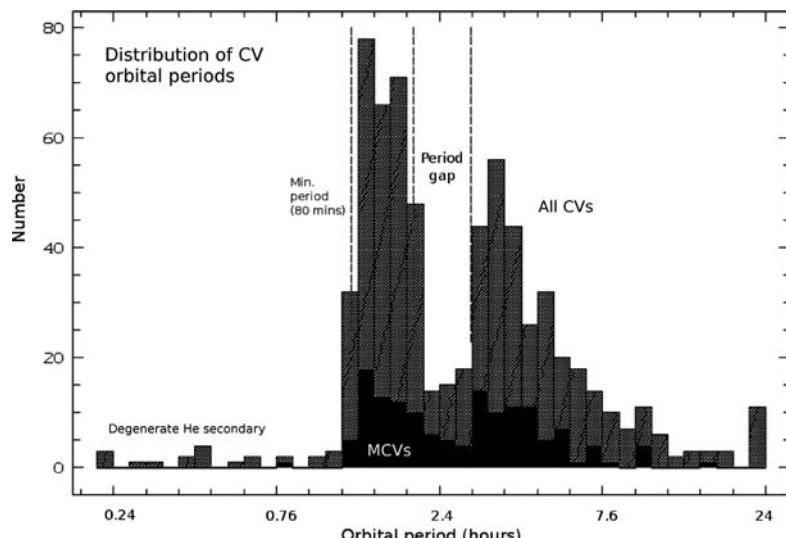
10.5.1 Evolution to short periods

Observationally, there are few CVs with orbital periods between 2 and 3 hours, a range which has come to be called the *period gap* (Fig. 10.11).

There are now more than 400 CVs with known orbital periods, so this diagram cannot be substantially distorted by selection effects. The gap must be due to a real physical effect. The secondary star is losing mass because it is filling its Roche lobe, the equipotential gravitational surface that surrounds the binary. It might be thought that it is the normal stellar evolution of the red star that causes it to expand and hence overflow its Roche lobe. For low-mass stars, however (as is the case for virtually all CVs), this cannot come close to accounting for the observed luminosities of CVs. Another mechanism is required that causes the Roche lobe to shrink.

Although still not well understood, this mechanism is believed to be magnetic braking. Just like the Sun, the CV's cool companion star will have a stellar wind. If the star also has a magnetic field, this wind will be forced to co-rotate out to

Fig 10.11 Distribution of orbital periods amongst cataclysmic variables. This clearly shows the much reduced number of CVs with periods in the range 2 to 3 hours (the period gap) and also the minimum orbital period of about 80 min. The only objects below the minimum period are double degenerates, in which the mass-losing star is also a white dwarf (adapted from a diagram by de Martino).



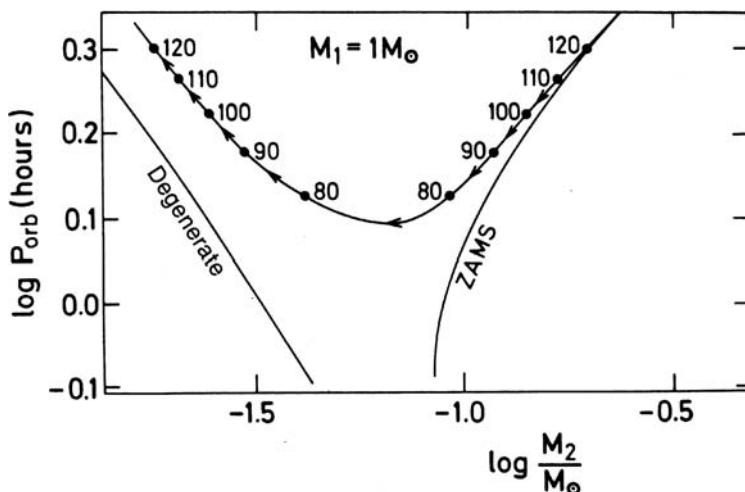


Fig 10.12 A minimum period for CVs. After exiting from the period gap, mass transfer from the secondary star will continue, and as long as the star behaves like a main sequence star, the secondary will shrink in size and the orbital period will also decrease. However, once the secondary approaches $0.1 M_{\odot}$ in size, it ceases to be able to burn hydrogen and leaves the main sequence. It becomes degenerate. Such a star behaves in the opposite way and increases in extent as it loses mass. The orbital period will then start to increase again after reaching a minimum at about 80 min (from work by Rappaport, Joss and Webbink, MIT).

some distance from the star. (In the Sun's case, this distance is about 100 solar radii, or about halfway to the Earth.) The energy required to do this will slow down the spin rate of the red star. (Imagine spinning around with arms outstretched. Then consider doing it whilst holding a long pole outstretched in each hand. The latter case will require a lot more effort and will slow down your rate of spin.) But the spin rate of the low-mass companion in close binaries is locked gravitationally and precisely to the orbital period, and so the effect of the drag imposed by the stellar wind will be to reduce the orbital angular momentum of the binary and hence cause the binary period to reduce. This is the mechanism that drives the mass transfer rate in CVs and LMXBs above the period gap.

However, at periods of about 3 hours, something happens (the braking effect may be reduced by the donor's wind being trapped by the white dwarf magnetosphere), and the mechanism is switched off. The donor star shrinks slightly and the mass transfer stops. Without its main power source, such a system would be much harder to find because it would no longer contain the accretion disc that provides its distinguishing characteristics of blue colour, variability, emission lines or X-ray emission. Orbital angular momentum is then lost by the process of gravitational radiation, and eventually, the Roche lobe shrinks again to catch up with the star's surface, and mass transfer begins again at a period of around 2 hours. The system is re-born and can be seen again as a CV.

How do such short-period systems evolve from this point? The binary is so small that gravitational radiation (GR) is now the main driving mechanism for the mass transfer as the period continues to reduce. However, the mass transfer rate under GR is much smaller than it was when the binary was above the period gap. If that is the case, though, how then is it possible to detect X-rays from them that are comparable to longer-period systems? The answer is that when mass transfer is restored on emerging from the period gap, there is a temporary enhancement in the mass loss as the secondary star settles down. This enhancement will eventually die away, and the source will become a much weaker X-ray emitter. Hence we would expect X-ray surveys to select those systems with larger mass transfer rates, which are those that have just emerged from the period gap; that is, most AM Her systems will have the same orbital period, which is what has been observed!

After this initial activity on emerging from the period gap, GR will act to reduce the orbital period as mass continues to be lost. The source will become much fainter and hence more difficult to detect. But the period does not decrease indefinitely. Models constructed by, amongst others, Saul Rappaport (MIT), Frank Verbunt (Utrecht) and Hans Ritter (Munich), show that the orbital period goes through a minimum at about 80 min (see Fig. 10.12).

This is because as the mass of the secondary star reduces, its nuclear evolution slows down, but

Table 10.3 | Properties of magnetic CVs (selected examples).

Star	P_{orb} (min)	P_{spin} (min)	V^1	d (pc)	F_X	kT_{BB} (eV)	B^2 (MG)
<i>Polars</i>							
EF Eri	81	Same	13.5	94		19	12
AR UMa	116	Same	15.1	88		23	230
AM Her	186	Same	12	75		~25	14
QQ Vul	226	Same	15.3		≥ 320	18–29	
<i>Intermediate polars</i>							
EX Hya	96	67	13.5	105	8.4		~1
V1223 Sgr	204	12.4	13.0	600	5.0		~1
TV Col	330	31.8	13.8	≥ 500	4.1		~1
GK Per	2874	5.9	13	340	3–15		~1

Note. All spectroscopically detected companion (mass donor) stars are cool and of low mass, with spectral types in the range M2–M7.

¹ Max brightness observed (most polars exhibit extended periods of quiescence when the mass transfer rate is greatly reduced, and their optical brightness decreases by ~2–5 mags).

² Derived from cyclotron harmonic structure in optical/infrared wide-wavelength spectroscopy. Estimates of magnetic field B exist for the second accreting pole in some of these systems (see Warner, 1995).

it is still losing mass. The star therefore departs from thermal equilibrium, and gravitational radiation now drives it to longer periods, but with no mass transferred. The ultimate fate of the system is a double degenerate binary consisting of a white dwarf and a very low mass black dwarf. Such systems would be extremely difficult to find.

10.6 Magnetic CVs

The CVs we have discussed so far are dominated in their observed behaviour by the accretion disc, the inner regions of which reach right down to the white dwarf itself. While these objects have benefitted from X-ray studies, they were all previously known. X-ray astronomy, however, has been responsible for the discovery of *magnetic CVs*, in which the magnetic field of the white dwarf is sufficiently powerful to take control of the accretion flow (see Box 10.3). Such systems either have a truncated disc (*intermediate polars*, also referred to as *DQ Her systems*) or no disc at all (*polars*, or *AM Her systems*). For polars, the magnetic field is sufficiently strong that the white dwarf rotation is locked to the orbital period, whereas in intermediate polars

(or IPs), where the field is slightly weaker, the white dwarf spins faster (see Table 10.3).

Box 10.3 Truncation of the disc in magnetic CVs

There will be a point at which the magnetic field of the white dwarf takes control of the material it is accreting from its companion. The radial distance at which this happens can be estimated by balancing the ram pressure $\sim \rho v^2$ of the accreting gas with the magnetic pressure $B^2/8\pi$ and is called the *magnetospheric radius*, R_M (see e.g. Frank *et al.*, 1992). This is a similar calculation to that performed for accreting neutron stars in LMXBs, where it is shown that $R_M \propto B^{4/7} \dot{M}^{-2/7}$, and hence, for short period binaries (where Kepler's third law tells us that the separation is small) and sufficiently powerful magnetic fields, R_M will exceed the binary separation, thereby making it impossible for an accretion disc to form. For a typical white dwarf ($0.85 M_\odot$; $0.01 R_\odot$) accreting at $\sim 2 \times 10^{-10} M_\odot \text{ yr}^{-1}$ in a 2-hour binary, this occurs for a magnetic field $B \sim 10^7 \text{ G}$.

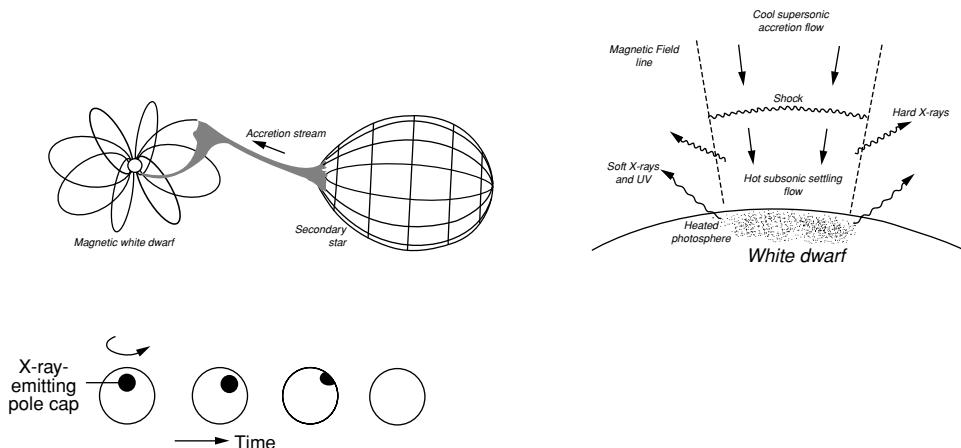


Fig 10.13 Schematic of a polar (AM Her system) in which the mass transfer from the secondary star is controlled by the powerful (tens of megagauss [MG]) magnetic field of the white dwarf. The accretion stream threads onto the field before it can form a disc, and it is then channelled onto the polar cap. With no disc, the matter is almost in free-fall, thereby creating a hard X-ray-emitting shock, which heats up the white dwarf surface. The modulation of the X-ray light curve on the orbital period is due to the occultation of the X-ray-emitting pole cap by the white dwarf itself (lower left).

10.6.1 AM Her systems (polars)

AM Her was actually detected by the Uhuru satellite as 4U1809+50 (and hence was a source of hard X-ray emission), but the optical identification with the CV itself was not made until SAS-3 located it to sufficient accuracy several years later, in the process showing that it was also a strong soft X-ray source and highly variable. However, its true nature (and the reason that it is the class prototype) was not revealed until the seminal work of Tapia (1977), who discovered extremely high circular polarisation which varied dramatically on a 3.09-hour period.

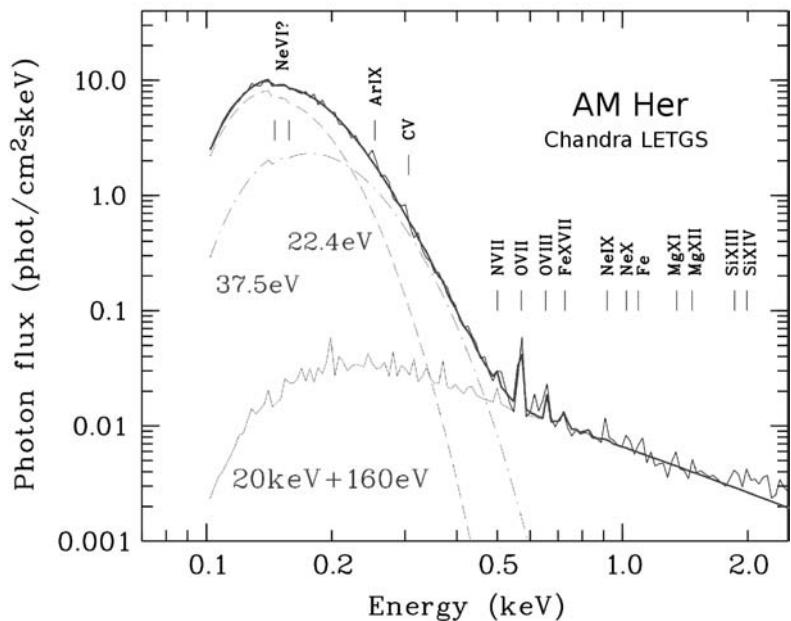
In polars (Fig. 10.13), the accretion stream from the companion is threaded by the intense magnetic field, which then channels the material into columns that lead to the white dwarf polar caps. The transferred material carries the orbital angular momentum of the companion, and this exerts a torque on the white dwarf, causing the rotation of the white dwarf to be locked to the orbital period. (There are a number of magnetic CVs that deviate from this rule, of which the most well known is V1500 Cyg, which was discovered as Nova Cyg 1975, a very bright naked-eye nova, and subsequently found to be a polar. Its rotation and orbital periods differ by just 1.8 per cent, but this is almost certainly a temporary effect caused by the instantaneous mass loss [$\sim 0.001 M_{\odot}$] that took

place during the nova ejection. The periods are expected to re-synchronise.)

The accretion column feeds matter in essentially free-fall onto the magnetic poles of the white dwarf (a very different process compared to accretion via a disc), reaching high velocities (typically $\sim 5000 \text{ km s}^{-1}$) before impacting the white dwarf. The flow of matter is halted in a strong shock above the surface (Fig. 10.13), where the kinetic energy is thermalised, corresponding to energies of $\sim 130 \text{ keV}$ for the protons. This energy is then radiated as hard X-rays by the electrons. These electrons are also spiralling in the magnetic field and so produce cyclotron radiation (predominantly in the optical and IR), which is strongly polarised (and was discovered by Tapia). The hard X-rays from the shock irradiate and heat up the (larger) white dwarf surface below, which then re-radiates this as an essentially blackbody soft X-ray component (another key characteristic of polars).

It is therefore expected that the soft X-ray flux observed from polars, L_{SX} , will be comparable to that of the hard X-ray (L_{HX}) and cyclotron (L_{cyc}) components from which it must be derived. However, and even in the light of better data and analyses (see e.g. Fig. 10.14), we still have $L_{BB} \geq 5(L_{HX} + L_{cyc})$, a discrepancy that became known as the soft X-ray puzzle. There are a variety of explanations (see e.g. Warner, 1995), but

Fig 10.14 Multi-component X-ray spectrum of AM Her obtained with the Chandra Low Energy Transmission Grating Spectrometer (LETGS). The very soft component requires two blackbodies (with $kT = 22$ and 37 eV) superposed on a much harder 20-keV continuum. There are also emission lines present due to ionised species which require a higher temperature (160 eV) coronal-type plasma (image courtesy of Klaus Reinsch and Klaus Beuermann).



it is generally accepted that the accretion flow from the donor star is not uniform but comes instead in ‘blobs’. If there were an accretion disc, this structure would be washed out completely, but in polars, these large blobs of material can penetrate the shock and the white dwarf surface before releasing their kinetic energy and subsequently heating up the white dwarf surface. The ‘puzzle’ is therefore explained as a reduction in the hard X-ray component. The X-ray spectrum is, in fact, more complex than this. For one thing, electron scattering can be important, and this makes the column’s X-ray spectrum seem harder when viewed side-on and softer at other angles (Rosen, 1992). Hard X-ray flares have also been seen high in the column (Cropper *et al.*, 2000).

10.6.1.1 Polars accreting on two poles

This simple polar model of accretion onto a single magnetic pole of the white dwarf can explain some of the basic X-ray light curves. As the white dwarf rotates, we see a varying fraction of the heated pole and shock region, thereby leading to a large modulation on the orbital/rotation period (Fig. 10.13).

However, the accretion geometry must necessarily involve two magnetic poles (Fig. 10.13), sometimes producing soft and hard X-ray light

curves that are almost in anti-phase with each other (Fig. 10.15). This means that the emission is a much more complex function of viewing angle and accretion column geometry (e.g. Cropper *et al.*, 2000). These light curves require that even the soft X-rays (which can only come from a tiny fraction, $\sim 10^{-4}$ to 10^{-5} of the white dwarf surface) must have some vertical extent, too.

The ROSAT all-sky survey (Trümper, 1993), because of its soft X-ray sensitivity, led to a large increase in the number of polars, the period distribution of which has a peak immediately below the period gap of Fig. 10.11 (only a handful are known above the gap). It is likely that this is related to a peak in the mass transfer rate during their current evolutionary phase since models predict substantially lower mass transfer rates during further evolution, and polars in this state will be much more difficult to discover. A notable feature of polars is that their optical brightness plummets almost instantly when the mass transfer rate drops as there is no disc present to act as a ‘buffer’. This provides one of the best means of directly studying the properties of the donor’s mass-loss mechanism. Hessman *et al.* (2000) have thereby exploited the 20-year light curve of AM Her assembled by the AAVSO to show that the average mass transfer rate is ~ 25 per cent of the maximum value,

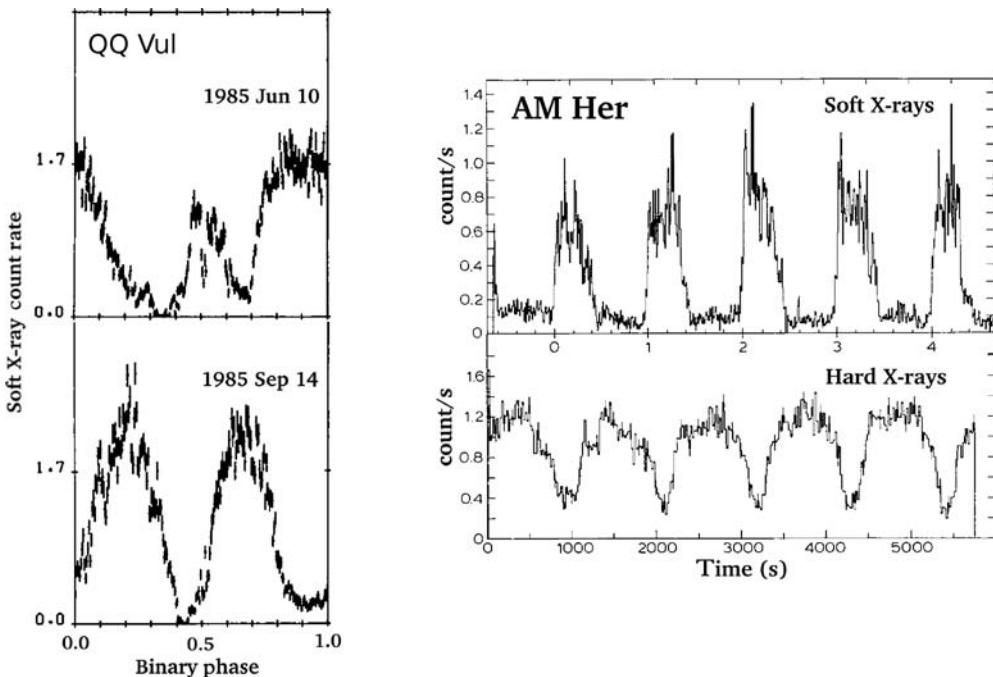


Fig 10.15 (left) Soft X-ray light curves of the polar QQ Vul as observed by EXOSAT on two occasions separated by 3 months in 1985 and plotted against orbital phase. The dominant emission on each occasion must be from completely different regions of the white dwarf, but note that they cannot be diametrically opposed. (right) Simultaneous soft and hard X-ray light curves of AM Her obtained with 1-min time resolution by EXOSAT using the low-energy (LE) and medium-energy (ME) telescopes. In this mode, AM Her is emitting in both bands from both magnetic poles (adapted from Osborne *et al.*, 1987 and Heise *et al.*, 1985).

and to map out the high and low state behaviour. These results have implications for the properties of dwarf novae, where the mass loss from the donor cannot be directly observed.

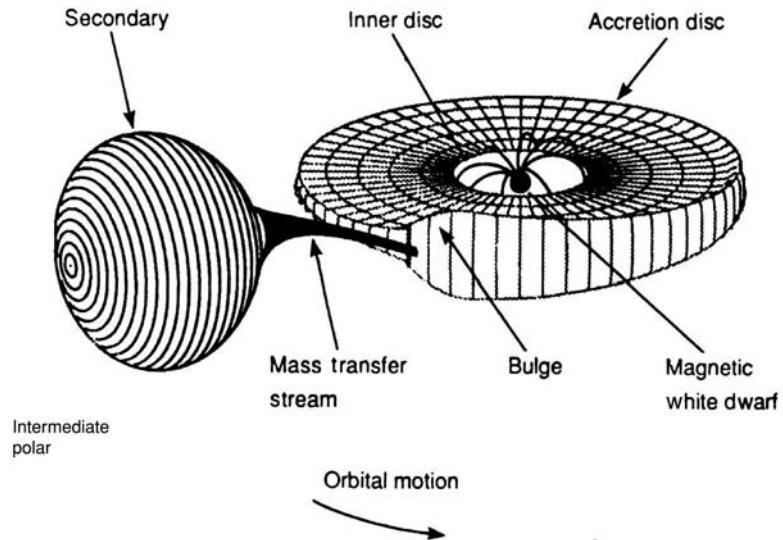
10.6.2 Intermediate Polars

The first of this sub-class discovered as an X-ray source was the CV EX Hya, which can be found in the Uhuru X-ray catalogue, but its unusual nature was not recognised at that time. The HEAO-1 X-ray instruments then located the source 2A0526-328, which was subsequently identified as TV Col (Charles *et al.*, 1979). Although recognised as CVs, their curious properties were mostly derived in the 1980s from optical studies (see the review by Patterson, 1994), followed by exploitation of EXOSAT's long satellite orbit (~ 4 days) to provide the first unbroken X-ray light curves lasting from many hours to days. These *intermediate polars* (IPs) exhibit some remarkable properties (for more details, see Buckley, 2000; Warner, 1995):

- X-ray pulsation in the range ~ 30 s to 1 hour due to the spinning magnetic white dwarf but which is *not* locked to the orbital period, as in polars
- optical pulsation on the spin period, together with orbital sidebands
- emission line variations on the orbital, spin and sideband periods
- a hard X-ray spectrum
- orbital modulations due to the observing geometry and absorption effects.

The clear spin modulation (in optical and X-ray) indicates that the magnetic white dwarf is controlling the accretion flow (as in polars), but the fact that this is not synchronised with the orbital period implies that the field is not strong enough to control the mass transfer from the donor as completely as it does in polars. Instead, a disc does form as in dwarf novae, but the inner regions of the disc are effectively truncated by the magnetic field. The presence of the disc makes it

Fig 10.16 The accretion geometry of an intermediate polar combines properties of both disc CVs and polars. The outer regions are remarkably similar to a ‘normal’ CV, but the inner regions are disrupted by the magnetic field of the white dwarf in much the same way as in polars (diagram by Keith Mason, MSSL).



more difficult to measure polarisation and hence determine magnetic field strengths, but this has now been done in a sufficient number of cases to support this inference (Table 10.3). This basic structure is shown schematically in Fig. 10.16, where material from the complete inner edge of the disc threads onto the magnetic field, creating a curtain of matter which can also modulate the observed signal.

Intermediate polars with shorter spin periods (approximately minutes) are sometimes referred to as ‘DQ Her systems’. DQ Her itself is not an X-ray source but is otherwise very similar to IPs. Consequently, here we include DQ Her and related systems in the IP classification (e.g. AE Aqr, which is a hard X-ray source, and GK Per [Fig. 10.17], with spin period immediately visible in the light curve). It is also clear when comparing the spin and orbital periods in magnetic cataclysmic variables (MCVs) in general (Fig. 10.18, top) that while a majority of IPs have spin periods around $0.1 P_{\text{orb}}$, there is a continuum of asynchronism values from $0.01 P_{\text{orb}}$ up to the polar line.

The truncated inner disc region of an IP is basically very similar to that in polars, but with the lower magnetic field producing accretion over a larger area of the white dwarf. This has been suggested as an explanation for the lack of soft component in IPs, but recently, several IPs have been revealed to have soft X-ray components (de

Martino *et al.*, 2008), which complicates the interpretation.

While a disc is unquestionably present in some IPs (e.g. EX Hya shows typical dwarf nova outbursts), it is not clear in others. However, the eclipsing systems have been essential for probing IP geometry, and time-resolved spectroscopy indicates that the stream does collide with an outer disc (see Hellier, 1993), as illustrated in Fig. 10.16. It is therefore not surprising that when the magnetic fields of polars and IPs are plotted together (Fig. 10.18, bottom), they appear as a continuum of values, with some overlap amongst MCVs at the high-field end of IPs and the low end of polars.

10.6.2.1 AE Aqr: A magnetic propellor

One of the IPs, AE Aqr, has quite unique properties compared to other systems. It does have an optical pulsation of 16.5 s, 33 s in X-rays (which is the rotation period) and a relatively long (9.9 hour) orbital period. Its brightness ($V \sim 11$) also meant that AE Aqr was an essential part of the development of the basic interacting binary model for CVs. Yet it now appears that, unlike in most other CVs discussed so far, most of the matter transferred from the secondary star is not actually accreted by the white dwarf. And its variability has a strange (flaring) signature of low amplitude (only ~ 1 mag, which has been known since 1929) and is very different from normal dwarf nova outbursts.

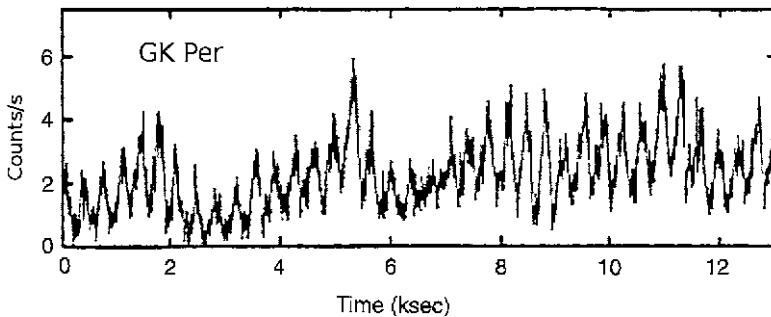


Fig 10.17 Hard X-ray light curve of the IP GK Per obtained with the EXOSAT observatory. The 6-min spin period of the white dwarf is clearly visible (adapted from Watson et al., 1985).

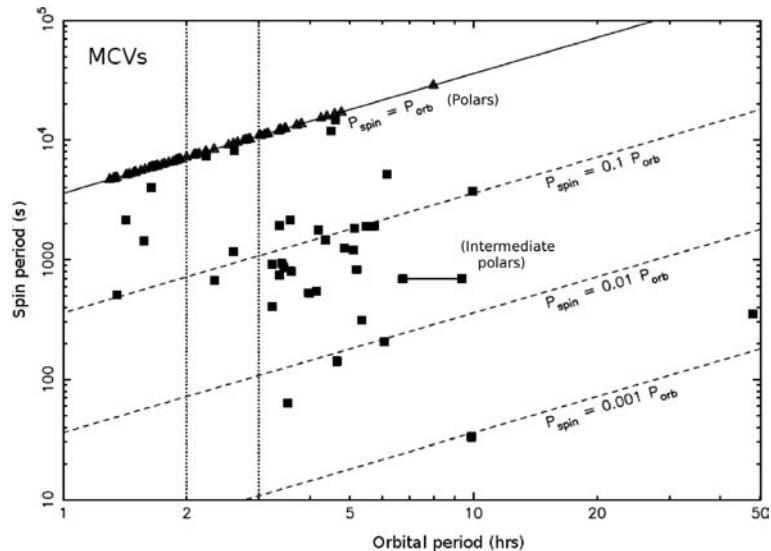
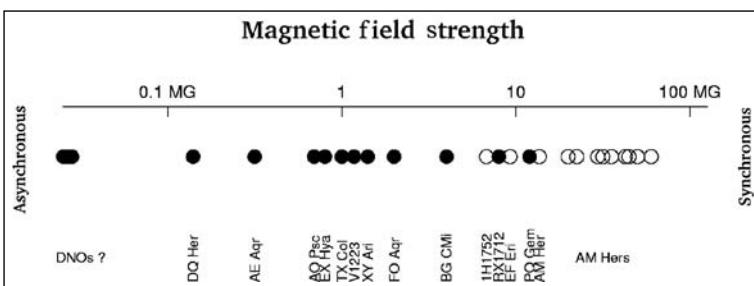


Fig 10.18 (top) Plot of spin period against orbital period for all magnetic CVs. The solid line corresponds to $P_{\text{spin}} = P_{\text{orb}}$, which includes all the polars (triangles). IPs are represented by squares, and the dotted lines give an indication of their degree of asynchronism.
(bottom) Distribution of magnetic field strengths for MCVs, showing that polars and IPs are not distinct but form a continuum (adapted from Hellier, 1996).

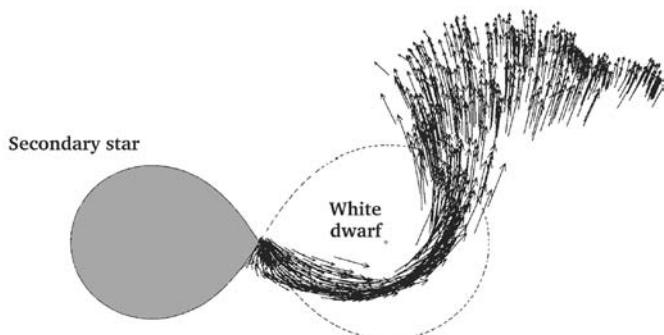


Consequently, AE Aqr has been subjected to intense multi-wavelength observing campaigns. These showed that the optical and UV flare together, as do the X-rays, but with a smaller amplitude. There are radio flares, too, but these are *not* associated with optical flares. More remarkably, AE Aqr is the only CV known to emit γ -rays during quiescent periods (they have been detected occasionally at extraordinarily high TeV energies). The flaring region must also be physically separate from the white dwarf's magnetic caps as the

pulsations were found not to be affected by the flares. Curiously, the H α (and other emission lines) do *not* follow the radial velocity curve of either the white dwarf or the companion star, nor is there an accretion disc present (the lines are not double peaked), even though the magnetic field is estimated to be weak enough to allow one to form. The total (bolometric) luminosity of AE Aqr (γ -ray, X-ray and optical) is $\leq 10^{33}$ erg s $^{-1}$, and yet the rate at which the white dwarf is spinning down implies a much higher energy loss of $> 10^{34}$ erg s $^{-1}$.

Fig 10.19 Schematic of the stream flow from the companion in the only γ -ray emitting CV, AE Aqr. The blobby structure of this flow leads to it being accelerated in the white dwarf's strong magnetic field and subsequently ejected from the system (adapted from Wynn et al., 1997).

Magnetic propellor model of AE Aqr



This ‘missing’ luminosity is explained by Wynn et al. (1997) with a ‘blobby’ stream flow from the companion. As the blobs approach the white dwarf, they are dragged and accelerated by the magnetic field to velocities that lead to their ejection from the system. They calculate that $\leq 10^{-4}$ of the material from the donor is actually accreted by the white dwarf, the rest being ejected (see Fig. 10.19), thereby carrying off kinetic energy, which we observe as the slowing of the spin period of the white dwarf. The flares arise when the blobs interact with the field closest to the white dwarf, creating shocks which account for the observed spectra.

Explaining the γ -rays and radio emission is more complicated, but we know that the white dwarf must have a magnetosphere because of the observed pulsations. Consequently, it has been suggested that protons accelerated in this magnetosphere have an impact on the blobs, thereby producing the observed γ -rays. And electrons accelerated in the same region produce the observed synchrotron radio emission.

10.7 X-ray spectroscopy of CVs

In general, the X-ray emission from CVs arises from regions close to the white dwarf and is due to the combined radiation from plasma at a range of temperatures (from the initial shock temperature to the white dwarf surface). What we ultimately see is controlled by the geometry of the accretion process, and this differs vastly between MCVs and non-magnetic systems. Koji Mukai and his colleagues

assembled comparable Chandra spectra of a number of different types of CVs to investigate their similarities and differences. The X-ray spectra fall into two classes:

1. These spectra (Fig. 10.20) all have strong H, He-like lines of O, Ne, Mg, Al, Si and S, plus the entire Fe L complex of lines from ions XVII to XXIV. These are well fitted by a multi-temperature plasma (plotted in the figure), exactly as would be expected to be emitted by gas that has been shocked to a very high temperature and then subsequently cooled. The model fitted here is the *cooling flow* spectrum first proposed for hot gas in clusters of galaxies. These CVs are all non-magnetic, apart from EX Hya.
2. The three IP spectra plotted here (Fig. 10.21) show a strong, hard power-law component, plus emission lines that are not well fit by a *cooling flow* as they show only FeXVII and not the remainder of the L-shell complex. This implies that the spectrum requires a *photoionisation* model, and this is a good fit.

The underlying physical difference between these two types of CVs is clearly related to their magnetic properties, and it is the mass transfer rate *per unit area*. As is clear in Fig. 10.13, the presence of the magnetic field focusses the accreting matter onto a small region of the white dwarf. The gas is shock heated to high temperatures and subsequently cooled by thermal radiation. However, the high density means that the majority of the X-rays will heat the gas in the flow prior to the shock, photo-ionising it and producing the spectra seen in type II.

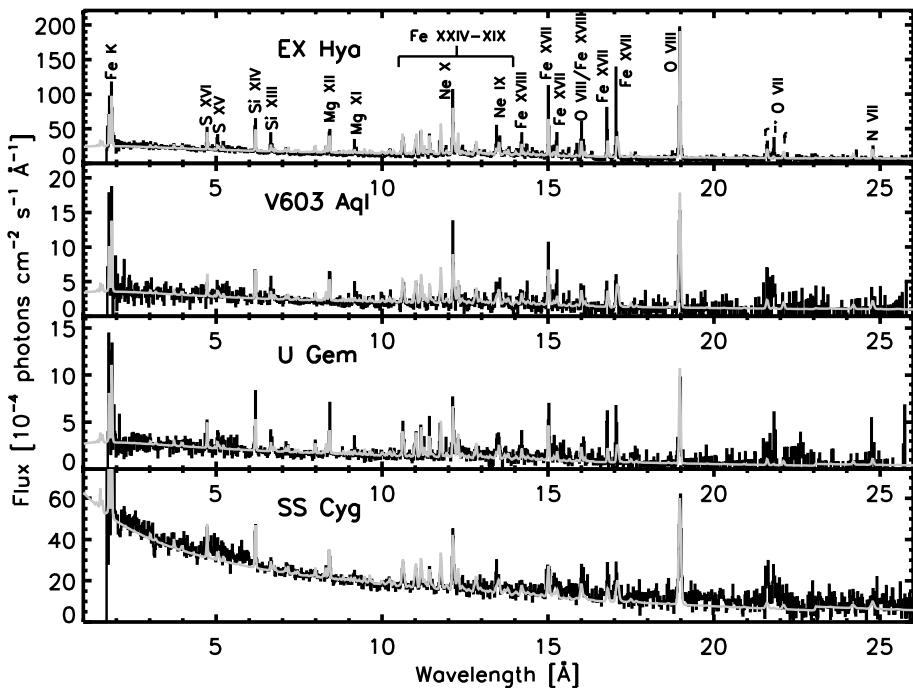


Fig 10.20 Chandra HETG spectra of four CVs, of which EX Hya is a magnetic system. They all display cooling flow-type spectra (model is shown as faint solid line) (from Mukai, 2003).

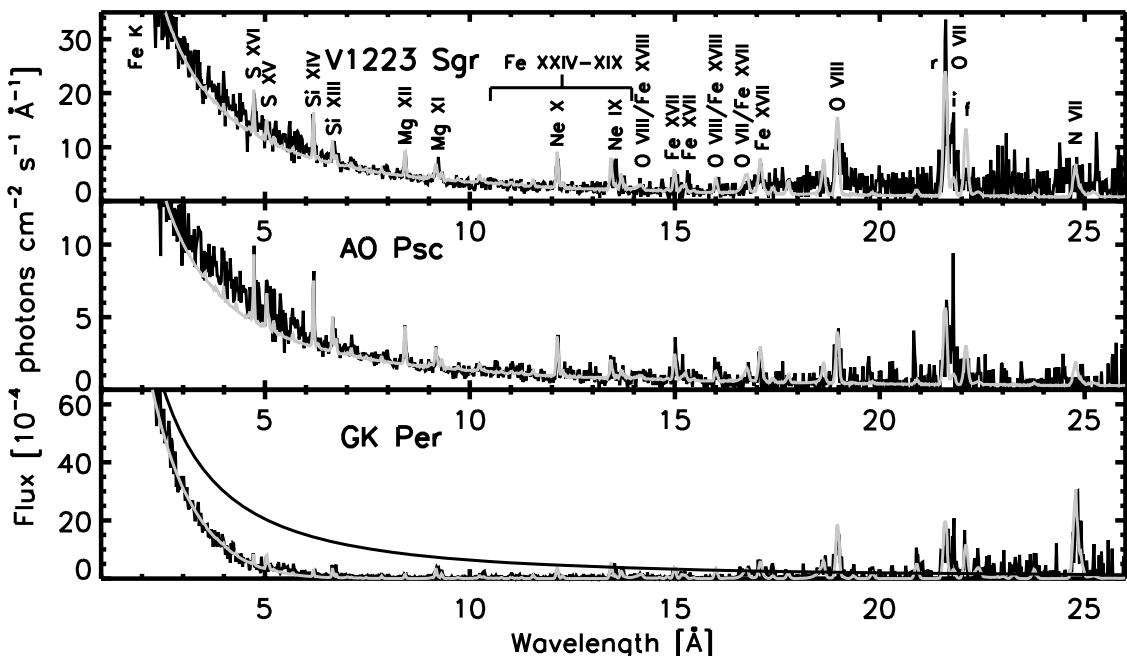


Fig 10.21 Chandra HETG spectra of three IPs, which are consistent with photoionisation models (shown as solid lines) (from Mukai, 2003).

If the mass transfer is more spread out (as in non-magnetic systems), then the X-rays from the post-shocked gas can radiate out the side, and we see this as the *cooling flow* spectra of type I. The reason that EX Hya is in this group, even though it is an IP, is because it has an unusually tall, low-density shock that also allows the cooling gas to radiate out of the side.

Whilst providing a broad interpretation of these X-ray spectra, it is likely that other CV properties (white dwarf mass, inclination, etc.) will also be important. These will require comparable quality spectra on a much wider range of CVs.

10.8 AM CVn systems

The variable AM CVn was discovered in 1967, and its 17-min modulation is now recognised as its orbital period, making it the prototype system of a class of *ultra-compact binaries*. Such an extreme system would have to consist of a white dwarf accreting from a helium-rich degenerate donor (a main sequence star could not fit inside the volume of such a short-period binary). The mass transfer would be driven by gravitational radiation, which would be significant in these cases.

Earlier it was pointed out that there is a minimum orbital period for CVs, as indicated in Fig. 10.12, so how are AM CVn systems formed? In some circumstances the donor star in a normal CV can produce sufficient mass transfer onto its white dwarf companion to initiate a second common-envelope phase. This will lead to the subsequent spiral-in of the white dwarf towards the core of the donor, resulting in a double white-dwarf system. If they are close enough, then this can result in one transferring matter onto the other and hence re-entering the CV phase, but at much shorter periods.

There are now a dozen or so AM CVn systems known, of which potentially the most extreme, RXJ0806.3+1527, was found by the ROSAT survey. It has a 5.4-min modulation of its X-ray flux (Fig. 10.22), which, if ultimately confirmed as the orbital period, will make it the shortest period binary known. The modulation is present in optical light, as well, with a phase lag that is consistent with X-ray heating of the donor. However,

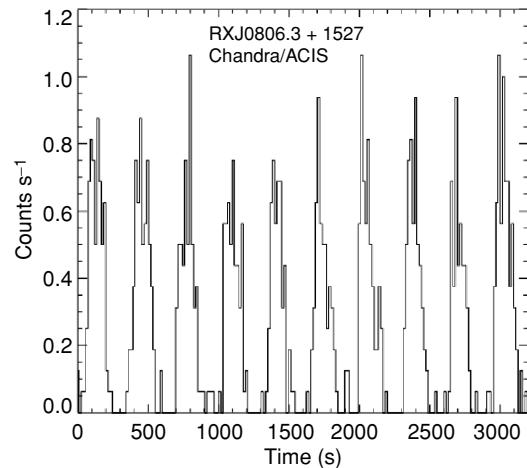


Fig 10.22 Chandra ACIS light curve of the putative AM CVn system, RXJ0806.3+1527, in which the 321-s modulation is clearly visible (from Strohmayer, 2005).

confirmation of this interpretation will require a radial velocity study, which is extremely challenging. Other interpretations are that the modulation may represent the spin period of an IP, but absence of any other modulation or the presence of polarisation does seem to favour the AM CVn interpretation. Such objects are of intense astrophysical interest because they are potentially strong sources of gravitational radiation, and their ultimate fate could be to end as type Ia supernova explosions.

10.9 Supersoft sources

The combination of soft X-ray sensitivity and direct imaging led to the discovery with Einstein of the very luminous ($\sim 10^{36-38} \text{ erg s}^{-1}$) but extremely soft (low temperature, $kT \sim 10-75 \text{ eV}$) class of objects that are now known as the *supersoft sources* (or SSS) (Long *et al.*, 1981). With such high luminosities ($> 10^3$ times that of typical CVs and dwarf novae in our own Galaxy), they were initially thought to be LMXBs. And by analogy with galactic black-hole systems such as LMC X-3 (on the basis of the softness of their spectra), it was even thought that SSS might harbour a black-hole compact object. However, BH LMXBs are also luminous in hard X-rays, and this is not

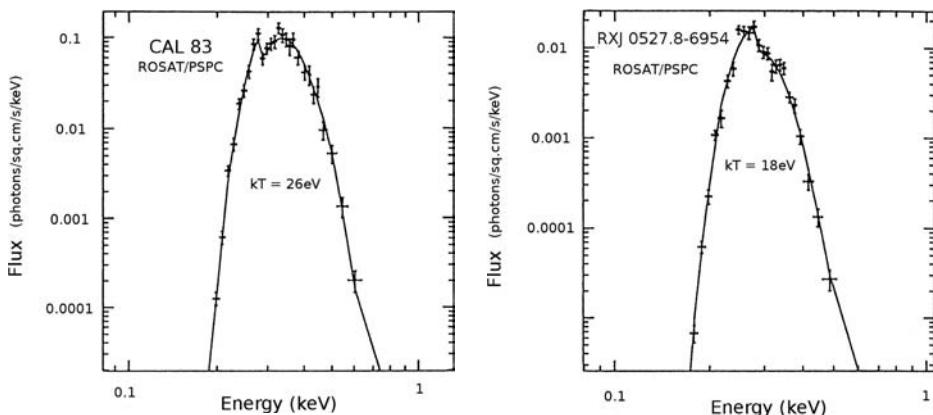


Fig 10.23 ROSAT PSPC spectra of the (left) SSS CAL83 and (right) RXJ0527.8-6954, together with their best-fitting blackbody spectra. All detected X-rays are at < 1 keV, and there is no hard component (from Greiner et al., 1991).

the case with SSS. Subsequent examples of SSS have been found in the SMC and M31 and more recently by Chandra and XMM in nearby spiral and elliptical galaxies. However, very few have been found in our Galaxy, as expected due to the effects of interstellar absorption on such low-energy X-ray sources. Hence the SSS can only readily be found and studied at high galactic latitudes using soft X-ray detectors such as the ROSAT PSPC (Fig. 10.23).

The breakthrough in understanding SSS came when van den Heuvel *et al.* (1992) noted that this combination of high luminosity and low temperature, if radiated as a blackbody, i.e. $L = 4\pi R^2 \sigma T^4$, can be rewritten

$$R = 10^4 \text{ km} (L / 10^{37} \text{ erg s}^{-1})^{0.5} (T / 3 \times 10^5 \text{ K}), \quad (10.3)$$

which has a size comparable to that of a white dwarf, not the much smaller neutron star or black-hole event horizon. However, to account for the very high luminosity requires an extremely high \dot{M} ($\sim 10^{-6} M_\odot \text{ yr}^{-1}$) if L_X is entirely accretion driven (because the potential well of a white dwarf is much shallower than that of a neutron star). Such an \dot{M} is 100 times that of a CV in outburst and would absorb all the soft X-ray flux.

However, there is another energy source locked within the accreting hydrogen, and that is fusion. In neutron star–black hole X-ray binaries, the accretion energy (per unit mass) is

15 times greater than that released by nuclear fusion (which can therefore be ignored under most circumstances), but in white dwarf systems, the situation is reversed, with fusion energy exceeding that of accretion by a factor 30. Hence the mass transfer rate required in the SSS is only $\sim 10^{-7} - 10^{-8} M_\odot \text{ yr}^{-1}$ if the accreted material is assumed to be undergoing stable nuclear hydrogen burning. It would essentially be a non-ejecting nova, as also suggested by Shara *et al.* (1977). Nevertheless, even this \dot{M} in the SSS is a factor of at least 100 times higher than that in normal CVs, implying that the SSS is in a special evolutionary state of some kind. Van den Heuvel and colleagues proposed that SSS have a binary mass ratio ~ 1 (i.e. a relatively massive secondary compared to normal CVs) which necessarily leads to unstable mass transfer onto the white dwarf. This phase is short-lived and will only last a few times 10^7 years for a $\sim 1 M_\odot$ companion, explaining the rarity of SSS.

This model is compelling and accounts for many of the SSS observed features in the long-period systems (> 0.5 days), but confirmation has proven elusive, largely because of the difficulty of obtaining accurate mass estimates for both the compact object and its companion. None have yet been obtained because no secondary star has yet been (optically) spectroscopically detected. The intrinsic spectrum has been completely overwhelmed by the steady-burning white dwarf and the surrounding inner disc. However, several

Fig 10.24 Optical spectrum obtained with the Anglo-Australian Telescope (AAT) of the SSS RXJ0513.9-6951, which is dominated by strong emission from hydrogen ($H\beta$) and ionised helium at $\lambda 4686$. Note the features marked S^+ and S^- on either side of the two strongest lines, which are believed to be red- and blue-shifted emission from HeII and $H\beta$ that is part of a bipolar outflow from the system (from Southwell et al. 1996).

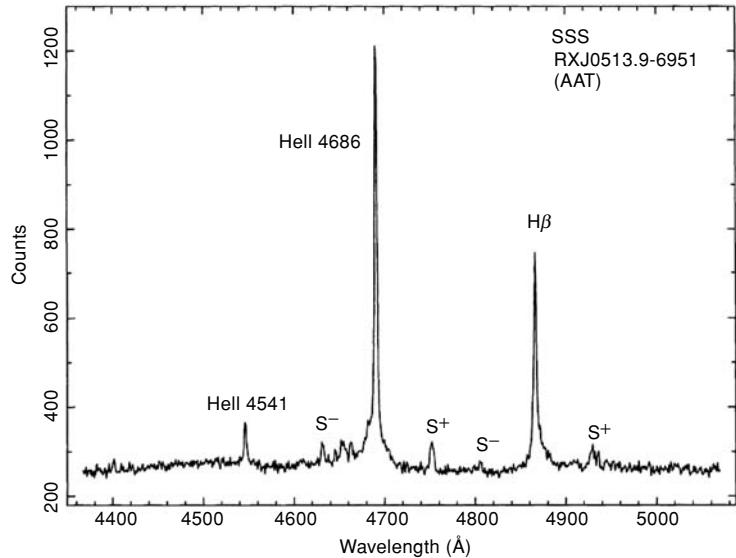
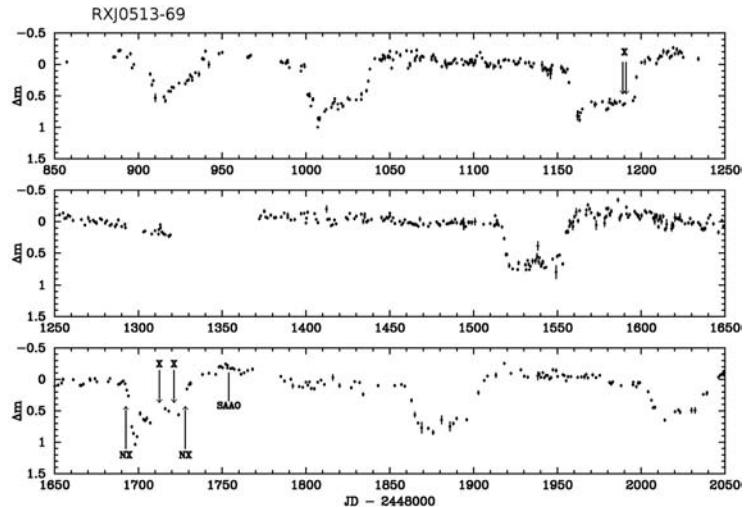


Fig 10.25 An almost 4-year optical light curve of the SSS RXJ0513.9-6951 obtained as part of the MACHO project. X-ray detections are marked with an 'X', non-detections as 'NX', showing that this SSS is only detectable during optical low states, making it effectively an 'SSS transient'.



pieces of indirect evidence strongly support the white dwarf model:

- Optical spectra of SSS have revealed evidence for high-velocity outflows (a bipolar outflow or disc wind) with velocity of $\sim 4500 \text{ km s}^{-1}$ (Fig. 10.24). Such outflows must, of course, exceed the escape velocity of the compact object, and in this case it is close to the escape velocity from a white dwarf.
- One of the SSS (RXJ0513.9-6951) shows transient X-ray behaviour that is anti-correlated with the optical flux (Fig. 10.25; McGowan et al., 2005). Night-to-night variations are punctuated

at irregular intervals by optical dips, and it is only at these times that the object emerges as an SSS. Detailed study of one of these intervals by XMM (Fig. 10.26) shows how the properties of the hot white dwarf surface evolve, and these are consistent with them being massive ($1.3-1.4 M_\odot$). It is exactly as expected in the van den Heuvel model, as an extremely high \dot{M} will inflate the white dwarf surface sufficiently to move its peak radiation into the EUV/UV. Occasional reductions in \dot{M} will lower the overall luminosity, but the spectral peak of the emission will move into soft X-rays as the effective radius of the white dwarf reduces. Recovery in

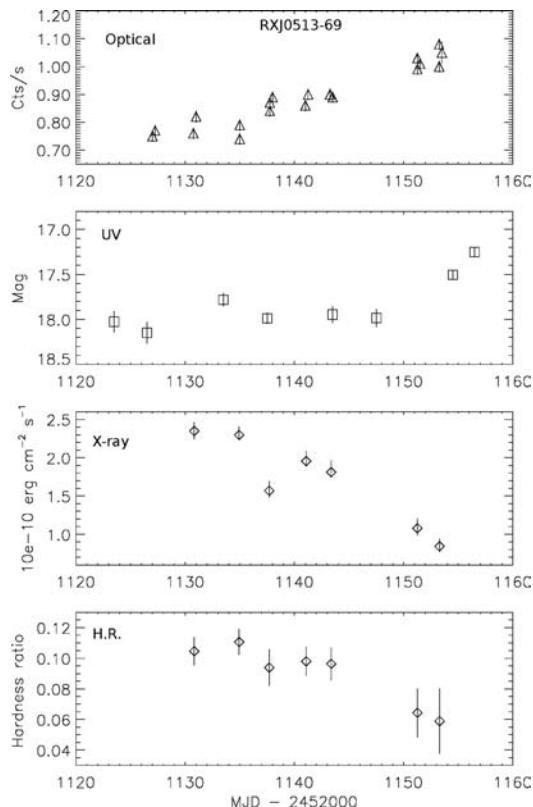


Fig 10.26 Detailed examination of an SSS-transient event in RXJ0513.9-6951 with XMM's X-ray and optical telescopes. The Optical Monitor (OM) produced the optical and UV light curves through the 'dip' (top), and EPIC-PN revealed the associated X-ray flux and spectral variability (bottom). Note that they are anti-correlated, with the X-ray flux diminishing and softening as the object brightens in the optical/UV.

the rate of \dot{M} (e.g. because of irradiation) will again inflate the white dwarf atmosphere, lowering the temperature and bringing the SSS phase to an end. CAL83 has shown evidence of similar behaviour.

- There is also excellent indirect evidence for the white dwarf nature of the compact object through ROSAT observations of the classical nova V1974 Cyg (Krautter *et al.*, 1996). While stable hydrogen burning does not occur in CVs, the accreted material on the white dwarf must eventually ignite in a thermonuclear runaway which we observe as a *nova* outburst and the ejection of a few $10^{-4} M_{\odot}$ shell at velocities of $\sim 5000 \text{ km s}^{-1}$. From a series of observations throughout the nova outburst of V1974 Cyg,

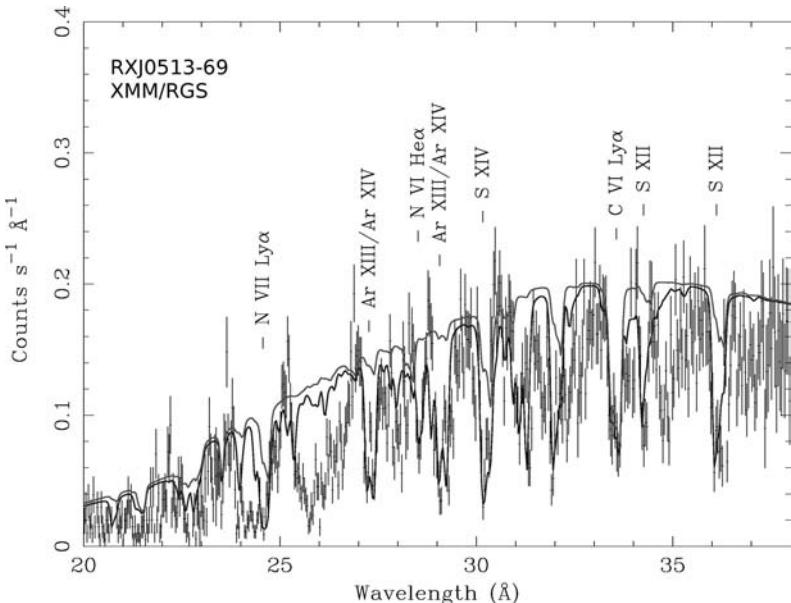
ROSAT revealed the brightest SSS yet, consistent with Eddington-limited emission from the white dwarf surface at temperatures of $\sim 20 \text{ eV}$. And since all novae have subsequently been shown to be CVs containing a white dwarf, this provides a compelling case for the basic SSS model.

Furthermore, grating spectroscopy with Chandra and XMM is revealing the rich and intricate absorption spectra of SSS (Fig. 10.27). The underlying continuum of RXJ0514-69 is represented as a 70-eV blackbody, but this is absorbed through substantial amounts of outflowing material (the identified features have velocity shifts). From this it is clear that the soft X-ray spectrum of this SSS is much more complex than expected. Enhanced nitrogen, argon and sulphur (relative to carbon) improve the fits to many of the lines, but other areas (e.g. around 26 Å) are likely blends of many features. More detailed analyses are now underway.

Very high accretion rates can also be sustained in another CV sub-class, the symbiotic binaries (see e.g. Kenyon, 1986). These are wide, very long period (hundreds of days) binaries in which a white dwarf is accreting from its red-giant companion. Mass transfer can be either by Roche-lobe overflow or via a dense stellar wind if the red giant is on the asymptotic giant branch (AGB), and examples of both types have been detected as SSS. However, the very high mass-loss rates implied by the dense winds observed ($\sim 10^{-5} M_{\odot} \text{ year}^{-1}$) mean that such a phase would be very short-lived.

The SSS are of considerable astrophysical interest because, as apparently massive white dwarfs, they are candidate objects to undergo accretion-induced collapse as their steady burning of accreted hydrogen provides a mechanism for their mass to grow. Whilst this might produce neutron stars formed via the accretion-induced collapse of a massive white dwarf, this is in contrast to eruptive variables, where all the accreted matter is ejected in nova explosions. Furthermore, the thermonuclear detonation of a mass-accreting CO white dwarf close to the Chandrasekhar limit ($\sim 1.4 M_{\odot}$) is the widely accepted explanation for type Ia supernovae (e.g. Branch *et al.*, 1995), and it is now considered possible for this to occur

Fig 10.27 XMM RGS spectrum of the SSS RXJ0513.9-6951 obtained during the transient event shown in Fig. 10.26. The underlying continuum is represented as a 70-eV blackbody, and this is absorbed through an outflowing, partially ionised wind with three phases. This absorption is modelled with (top) solar (dotted line) and (bottom) modified (solid line) abundances (from McGowan *et al.*, 2005).



in lower-mass white dwarfs via a double detonation involving the large He shell (produced by burning the accreted H) (see Ruiz-Lapuente *et al.*, 1995). The SSS are thus candidates for implication in all these processes. For a detailed review of the SSS, see Kahabka and ven den Heuvel (2006).

10.10 | Summary

Cataclysmic variables are white-dwarf interacting binaries with many similarities to low-mass X-ray binaries, except that they are far more numerous than their neutron star or black-hole counterparts. Consequently, the great majority of CVs which have been studied in detail are relatively nearby. They also provide ideal laboratories for examining the accretion process and the properties of accretion disks around both magnetic and non-magnetic white dwarfs. In the latter it is usual for the accretion disc's intrinsic emission to dominate, thereby revealing properties of discs that have implications for discs in a much wider range of astrophysical environments. The relatively high temperatures achieved in the inner disc and close to the accreting white dwarf mean

that space observations, particularly in the UV and X-ray regions, have proved crucial in unravelling the interaction of the accretion flow with the disc and the white dwarf.

References

- Branch, D., *et al.* (1995) In search of the progenitors of type IA supernovae, *PASP* **107**, 1019.
- Buckley, D. A. H. (2000) On the power in intermediate polars, *New Astron Rev* **44**, 63.
- Charles, P. A., *et al.* (1979) 2A 0526-328 – an X-ray-emitting cataclysmic variable, *ApJ* **231**, L131.
- Córdova, F. A. & K. O. Mason (1984) X-ray observations of a large sample of cataclysmic variable stars using the Einstein Observatory, *MNRAS* **206**, 879.
- Cropper, M., K. Wu & G. Ramsay (2000) The emission from post-shock flows in mCVs, *New Astron Rev* **44**, 57.
- De Martino, D., *et al.* (2008) GALEX ultraviolet photometry of NGC 2420: Searching for WDs, *Mem SAI* **79**, 246.
- Eracleous, M. & K. Horne (1996) The speedy magnetic propeller in the Cataclysmic Variable AE Aquarii, *ApJ* **471**, 427.
- Frank J., A. King & D. Raine (1992) *Accretion Power in Astrophysics*, Cambridge: Cambridge University Press.

- Greiner, J., G. Hasinger & P. Kahabka (1991) ROSAT observation of two supersoft sources in the Large Magellanic Cloud, *A&A* **246**, L17.
- Heise, J., et al. (1978) Detection of both soft and hard X-ray emission from SS Cygni with ANS, *A&A* **63**, L1.
- Heise, J., et al. (1985) An X-ray study of AM Herculis. I – Discovery of a new mode of soft X-ray emission, *A&A* **148**, L14.
- Hellier, C. (1993) The accretion stream in intermediate polar binaries, *PASP* **105**, 966.
- Hellier, C. (1996) The intermediate polars, in *Cataclysmic Variables and Related Objects*, 143, edited by A. Evans & J. H. Wood, New York: Kluwer.
- Hessman, F. V., B. T. Gänsicke & J. A. Mattei (2000) The history and source of mass-transfer variations in AM Herculis, *A&A* **361**, 952.
- Kahabka, P. & E. P. J. van den Heuvel (2006) Supersoft Sources, in *Compact Stellar X-ray Sources*, 461, edited by M. van der Klis and E. P. J. van den Heuvel, Cambridge: Cambridge University Press.
- Kenyon, S. (1986) *The Symbiotic Stars*, Cambridge: Cambridge University Press.
- Kiplinger, A. L., E. M. Sion & P. Szkody (1991) A study of the ultraviolet evolution of U Geminorum between outbursts, *ApJ* **366**, 569.
- Knigge, C., et al. (1998) Recovery of 29 second oscillations in HST eclipse Observations of the CV UX UMa, *ApJ* **499**, 429.
- Krautter, J., et al. (1996) ROSAT X-ray observations of Nova V1974 Cygni: The rise and fall of the brightest supersoft X-ray source, *ApJ* **456**, 788.
- Krzeminski, W. & K. Serkowski (1977) Extremely high circular polarization of AN UMa, *ApJ* **216**, L45.
- Ia Dous, C. (1989) Synthetic optical and ultraviolet spectra of stationary accretion disks, *A&A* **211**, 131.
- Lasota, J. P. (2001) The disc instability model of dwarf novae and low-mass X-ray binary transients, *New Astron Rev* **45**, 449.
- Long, K. S., D. J. Helfand & D. A. Grabelsky (1981) A soft X-ray study of the Large Magellanic Cloud, *ApJ* **248**, 925.
- Marsh, T. R. (2001) Observations of cataclysmic variable and double degenerate stars, *Lect Notes Phys* **563**, 151.
- McGowan, K. E., et al. (2005) XMM spectroscopy of the transient supersoft source RXJ0513.9 – 6951: Probing the dynamic white dwarf photosphere, *MNRAS* **364**, 462.
- Mukai, K. (2003) X-ray spectroscopy of cataclysmic variables, *Adv Space Res* **32**, 2067.
- Osborne, J. P., et al. (1987) A new soft X-ray mode in the AM Herculis object E2003 + 225, *ApJ* **315**, L123.
- Patterson, J. (1994) The DQ Herculis stars, *PASP* **106**, 209.
- Patterson, J. & C. Price (1981) Spectrophotometry of AM Herculis at minimum, *PASP* **93**, 71.
- Patterson, J. & J. C. Raymond (1985) X-ray emission from cataclysmic variables with accretion disks. I – Hard X-rays. II – EUV/soft X-ray radiation, *ApJ* **292**, 535.
- Pringle, J. E. (1981) Accretion discs in astrophysics, *ARA&A* **19**, 137.
- Proga, D. (2003) On resonance-line profiles predicted by radiation-driven disk-wind models, *ApJ* **592**, L9.
- Proga, D., J. M. Stone & J. E. Drew (1999) Line-driven disc wind models with an improved line force, *MNRAS* **310**, 476.
- Ramsay, G., et al. (2002) Optical/infrared spectroscopy and photometry of the short-period binary RX J1914+24, *MNRAS* **333**, 575.
- Rappaport, S., et al. (1974) Possible detection of very soft X-rays from SS Cygni, *ApJ* **187**, L5.
- Rosen, S. R. (1992) The role of electron scattering in the X-ray rotational light curves of Intermediate Polars, *MNRAS* **254**, 493.
- Ruiz-Lapuente, P., A. Burkert & R. Canal (1995) Type IA supernovae scenarios and the Hubble sequence, *ApJ* **447**, L69.
- Shakura, N. I. & R. A. Sunyaev (1973) Black holes in binary systems: Observational appearance, *A&A* **24**, 337.
- Shara, M. M., D. Prialnik, & G. Shaviv (1977) Non-ejecting novae as EUV sources, *A&A* **61**, 363.
- Southwell, K. A. et al. (1996) The nature of the supersoft X-ray source RX J0513-69, *ApJ* **470**, 1065.
- Strohmayer, T. E. (2005) Precision X-ray timing of RX J0806.3+1527 with Chandra: Evidence for gravitational radiation from an ultracompact binary, *ApJ* **627**, 920.
- Tapia, S. (1977) Discovery of a magnetic compact star in the AM Herculis/3U 1809+50 system, *ApJ* **212**, L125.
- Trümper, J. (1993) ROSAT – A new look at the X-ray sky, *Science* **260**, 1769.
- van den Heuvel, E. P. J., et al. (1992) Accreting white dwarf models for CAL 83, CAL 87 and other ultrasoft X-ray sources in the LMC, *A&A* **262**, 97.
- Warner, B. (1972) Observations of rapid blue variables – VII. EX Hyd, *MNRAS* **158**, 425.
- Warner, B. (1995) *Cataclysmic Variable Stars*, Cambridge: Cambridge University Press.

- Warner, B. & R. E. Nather (1972) Observations of rapid blue variables – XII. UX UMa, *MNRAS* **159**, 429.
- Watson, M. G., A. R. King & J. P. Osborne (1985) The old nova GK Per – Discovery of the X-ray pulse period, *MNRAS* **212**, 917.
- Wheatley, P. J., C. W. Mauche & J. A. Mattei (2000) RXTE, EUVE and optical observations of SS Cyg in outburst, *New Astron Rev* **44**, 33.
- Wynn, G. A., A. R. King & K. Horne (1997) A magnetic propeller in the cataclysmic variable AE Aquarii, *MNRAS* **286**, 436.

Chapter 11

X-ray binaries

11.1 Introduction

11.1.1 The discovery of binary behaviour

The very existence of the bright cosmic X-ray sources discovered in the 1960s represented an exciting and challenging astrophysical problem. No physical process known then was capable of generating the enormous X-ray luminosities observed. The subsequent optical identifications of Sco X-1 and Cyg X-2 stimulated theorists and observers alike to learn more about these new ‘X-ray stars’. Why were these extremely powerful X-ray sources associated with such apparently unremarkable optical objects (see Chapter 1)? They were rather faint (13th to 15th magnitude) and did not stand out on optical photographs. However, the optical spectrum of Sco X-1 had similarities with the cataclysmic variables that were being intensively monitored by amateur groups and had been shown, a few years earlier, to be interacting binary systems (see Chapter 10).

As shown in Fig. 11.1, Sco X-1 displayed a smooth blue continuum with superposed emission lines of hydrogen and ionised helium. The absence of absorption features, as in normal stellar spectra, indicated that little or none of the light was coming from a main sequence star. The presence of ionised helium indicated that the source of excitation of the lines was very hot and very likely to be connected with the X-rays. However, despite many observing campaigns dedicated to Sco X-1, which revealed substantial variability on all timescales, no indication of binary

behaviour was found. The same was true for Cyg X-2.

Nevertheless, there was a great deal of theoretical speculation about the nature of the energy source in Sco X-1, and this centred around what had previously been considered to be very exotic objects: white dwarfs, neutron stars and black holes. Indeed, at this time (1966), pulsars had still not been discovered, and neutron stars existed only as a theoretical concept.

Our fundamental understanding of the nature of the bright X-ray stars came from the Uhuru satellite in 1971. Shortly after launch, it discovered two sources which were pulsating regularly with precise periods (Giacconi *et al.*, 1971; Tananbaum *et al.*, 1972). When observed over much longer time intervals (Figs. 11.2 and 11.3), both sources showed Doppler shifts and X-ray eclipses (Schreier *et al.*, 1972). It was recognised immediately that they must both be close binary systems in which the X-ray emitting object was a rapidly spinning neutron star (it could not be a white dwarf because the period was so short that a white dwarf would break up due to centrifugal force). The short binary period indicated that the two components were probably close enough together to be interacting, i.e. likely to be exchanging mass (see Box 11.1). The X-ray energy source was then gravity. Matter is transferred from one component onto the compact object, thereby releasing a large amount of gravitational energy in the form of X-radiation. (That matter heats up in falling in this way can be nicely demonstrated by a simple experiment where the temperature of

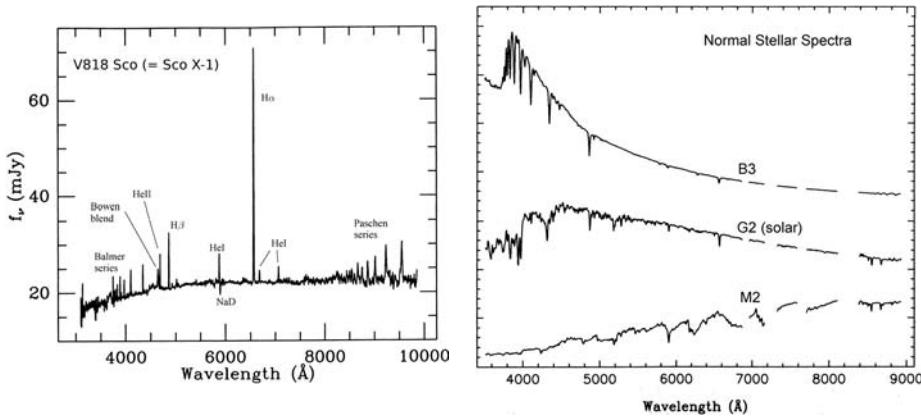


Fig 11.1 (left) Optical spectrum of Sco X-1 compared with (right) spectra of ‘normal’ (i.e. non-interacting) stars covering a range of spectral types from hot, early-type (B3) through solar (G2) to very cool (M2). Note that the normal stars all display easily recognisable absorption features due to hydrogen (B3) through calcium, sodium and magnesium (solar) to molecular TiO (M2), whereas none are present in Sco X-1 (apart from interstellar sodium). Instead, Sco X-1 exhibits strong emission lines due to hydrogen (especially H α , β , the rest of the Balmer series and the Paschen lines beyond 8500 Å) and ionised carbon, nitrogen (the ‘Bowen blend’) and helium (between 4650 and 4700 Å). Hell $\lambda 4686$ especially indicates the presence of very hot material. Sco X-1 also has a blue continuum (it would appear bluer here if it were not for interstellar absorption) which has a different shape to that of any of the normal stars. Spectra from Schachter et al., 1989 (Sco X-1) and Silva & Cornell, 1992 (normal stellar spectra).

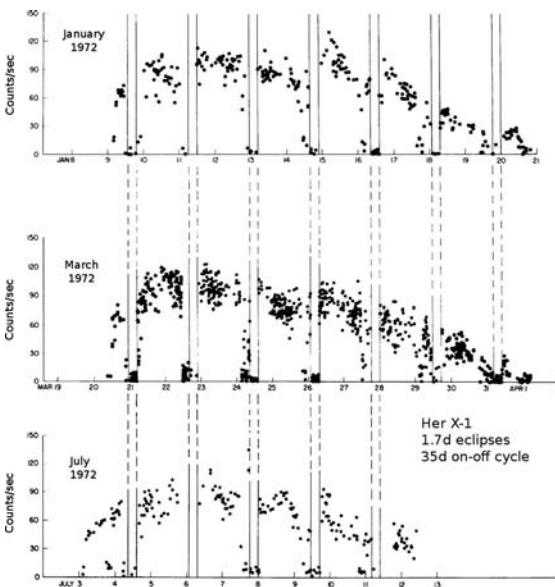


Fig 11.2 Uhuru X-ray light curve of Her X-1 which clearly demonstrated that it was an eclipsing binary, in which the X-ray source regularly disappears behind its companion.

water is determined at the top and bottom of a waterfall.)

It was quickly realised that there were two basic forms that binary systems could take in

which mass could be transferred from an (almost) normal star onto its orbiting compact object (see Fig. 11.4):

1. *High-mass X-ray binaries (HMXBs)*. In these binaries the material for transfer onto the compact object is provided by a powerful stellar wind from an early-type, massive star similar to those in the OB associations discussed in Chapter 7. Optically we see only the early-type star because it is so luminous. Cen X-3 and Cyg X-1 are both in this class.

2. *Low-mass X-ray binaries (LMXBs)*. If the mass-losing star is of low mass, then it will not have a strong stellar wind and hence cannot power the X-ray source by the same mechanism as in HMXBs. In LMXBs the companion star has evolved to fill its Roche lobe (surface of equal gravity) and is transferring material through the *inner Lagrangian point*, L_1 , onto the compact object (see Fig. 11.4 and Box 11.2). Such mass transfer is only stable if the donor star is less massive than the compact object. For neutron star systems, this means that the donor stars must be less massive than the Sun (sometimes much less massive), and hence LMXBs will be optically faint, with the donor star usually undetectable. Sco X-1 is in fact an LMXB, as is Her X-1.

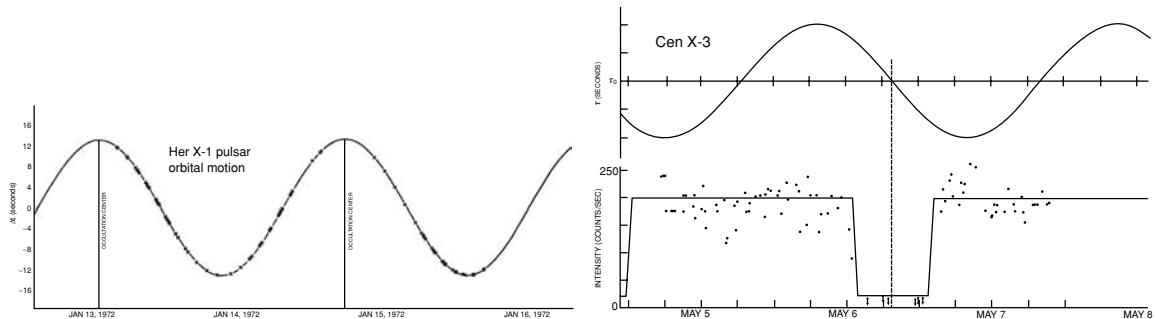


Fig 11.3 The regular changes in the spin periods of Her X-1 and Cen X-3 were observed with Uhuru, as the X-ray pulsars orbited their companion stars every 1.7 days and 2.09 days, respectively. These pulse arrival time curves give the velocity of the pulsars' motion and, combined with optical observations of the donors, allow the stars' masses to be determined (see Box 11.4).

These X-ray pulsars (Cen X-3 and Her X-1) became cornerstones in the study of X-ray binaries because, as we shall see, they provided detailed information about the material near the compact object itself. In addition, even though the sky positions of X-ray sources determined by Uhuru were relatively poor (especially compared to those produced subsequently by SAS-3 and Einstein), and there were consequently many stars in the error boxes (sky charts of the locations), the X-ray pulsations and

binary periods provided signatures with which to hunt for the optical counterpart. In fact, this proved invaluable in locating Cen X-3 as the visible star was only confirmed by the discovery of an optical modulation at the binary period.

These discoveries transformed and stimulated the field of X-ray astronomy into a period of exciting growth that has been maintained until the present day, turning it into a major branch of observational astronomy.

Box 11.1 | Interacting binaries

Once the orbital period, P , of a binary system is known, Kepler's third law can be applied to estimate the separation, a , of the two masses:

$$a = 4.2(M_1 + M_2)^{1/3}P^{2/3}, \quad (11.1)$$

where a is measured in solar radii, the masses of the two stars are in M_\odot and P is in days. Since most X-ray binaries have short periods (<20 days) and $M_1 + M_2$ is in the range 1–10 M_\odot , then the separation of the two stars will be not much larger than the size of the mass-losing star itself. In the case of HZ Her/Her X-1, which is eclipsing, we have a direct measure of the radius of the companion star, and the binary separation is only about twice this value.

The observed X-ray pulsations, which are due to the spin of the compact object (nothing else could produce such an accurate clock), show that the compact object cannot be a white dwarf and

hence must be a neutron star. Obviously the star is not spinning so fast that it will break up, so the centrifugal force felt by a mass m at the surface of the spinning star must be less than the force of gravity.

This gives $mv^2/R = G M m/R^2$, and since $v = 2\pi R/P$, R may be expressed in terms of M and P for the star to remain stable:

$$R < (G M P^2/4\pi^2)^{1/3}. \quad (11.2)$$

If the spin period is 1 s and M is 1 solar mass, then $R = 1500$ km is the maximum radius of the compact object, for at this size gravity is only just balancing the centrifugal force. This is much too small to accommodate a white dwarf, which has a typical radius of 10 000 km. The object must therefore be a neutron star, which has a typical radius of 10–15 km, well within the stability criterion evaluated here.

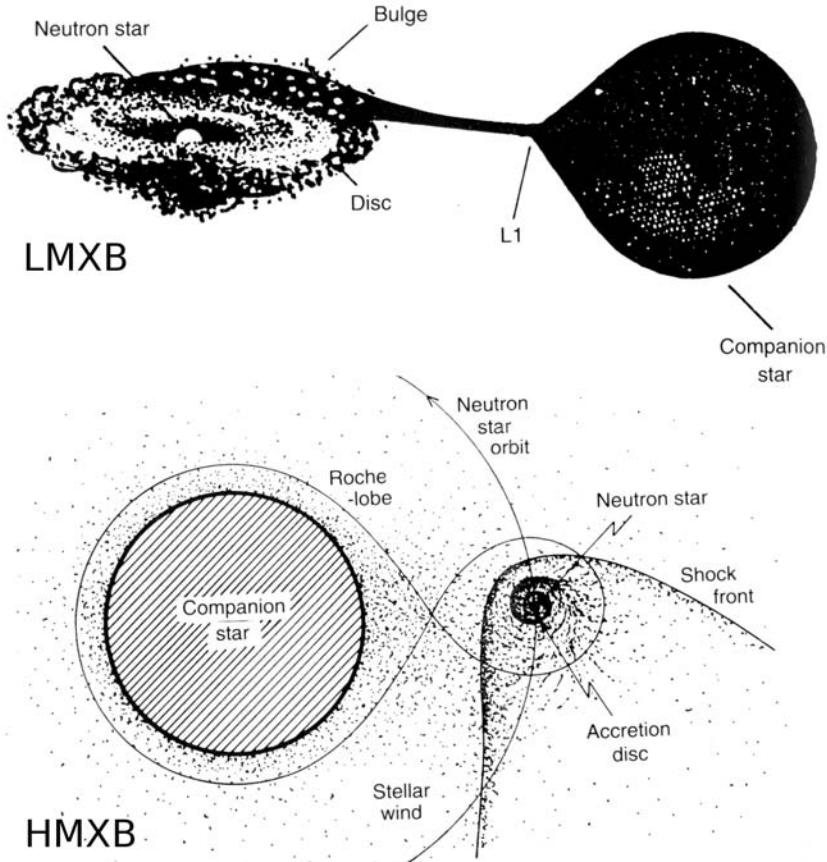


Fig 11.4 Artist's conception of the two principal mechanisms by which matter is transferred onto a compact object in a binary system. (top) A low-mass star has evolved and is losing mass to its degenerate companion because it fills its Roche lobe and mass flows through the L_1 point. It is the presence of the compact object that distorts the mass-losing star into this shape (rather like a pear), the vertex of which is the inner Lagrangian point, or L_1 . Matter flowing out of the star forms a stream that impacts the accretion disc (creating a bulge or thickened region) and, by viscous forces, is gradually accreted onto the compact object where the X-rays are generated. The view is for binary phase 0.25. (bottom) The compact object is orbiting a massive star which has a very powerful stellar wind – so powerful that there is sufficient material being lost in all directions for the compact object to accrete and produce copious X-rays as it ploughs through this wind, thereby creating a comet-shaped shock front. The accretion disc is small, and so fluctuations in wind density are immediately evident in the X-ray flux and pulsar frequency (courtesy EXOSAT Observatory, ESA).

11.1.2 Field guide to X-ray double stars

This brief and historical introduction to X-ray binaries hides what is now a bewildering variety of objects and behaviour patterns. At times it seems there are as many classes as there are objects to put in them! Indeed, we shudder to think how many times different objects' behaviour has been described as *unique!* We will describe the most important features of these binaries to indicate

their relationships to one another and to show how they reached their present state. It is important to remind the reader, though, that this continues to be a very active field, and the detailed evolution of binary stars as they transfer material between them is sometimes a subject of intense debate.

Basically, many of the properties of a binary depend on the nature of the compact object – is it

Box 11.2 | Mass transfer through the Lagrangian point

In interacting binaries, one of the pair of stars must be transferring material onto the other (which in a luminous X-ray source must be a compact object). This transfer is subject to the combined gravitational potential of the pair of stars, and this is plotted in Fig. 11.5. Close to the centre of each star, the potentials are almost spherical. However, as we approach the centre of mass of the binary, the potentials are distorted into a ‘pear’ shape, with the critical equipotential being that which includes the L_1 (or *inner Lagrangian*) point at which the net force is zero. This is the *Roche lobe* of each star.

The potential wells in which matter associated with both stars are constrained to move are shown by the one-dimensional plot of the potential along the x axis, which is defined by

$$V(x, 0) = -G \left(\frac{M_1}{x_1 - x} + \frac{M_2}{x_2 + x} \right) - \frac{1}{2} \omega^2 x^2, \quad (11.3)$$

where the angular frequency ω is $2\pi/P$, and from Kepler’s third law, $\omega^2 = G(M_1 + M_2)/a^3$. The additional term is due to the binary’s circular motion and gives rise to the other Lagrangian points shown (see Frank *et al.*, 2002, for further details). The mass-losing star, M_2 , fills its Roche lobe, and since the potential of L_1 is lower than L_3 , as the star expands, matter will be transferred through L_1 (a saddle point) into the potential well of M_1 .

There are two useful (but non-physical) approximations that give the equivalent volume (of radius R_2) of the Roche lobe of the mass donor as a function of the binary mass ratio $q (= M_2/M_1)$ and separation a :

$$\frac{R_2}{a} = 0.46 \left(\frac{q}{1+q} \right)^{1/3} \quad (\text{valid for } q < 1) \quad (11.4)$$

$$\frac{R_2}{a} = 0.38 + 0.2 \log q \quad (\text{valid for } 0.5 < q < 20). \quad (11.5)$$

These are due to Paczynski (1971) and are accurate to a few per cent.

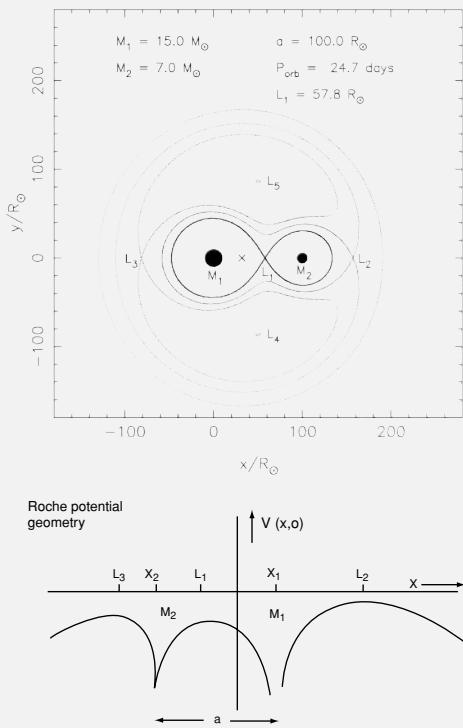


Fig 11.5 Lines of constant gravitational potential for a binary with the parameters shown. The *equipotential surface* that envelopes both stars and intersects at a point (L_1) is the critical surface and defines the Roche lobe of each star. Including the forces due to the rotation of the binary, there are five points (L_1 to L_5) where the forces cancel out, and these are known as the *Lagrangian points*. The shape of the equipotential surface is plotted along the x axis. Since L_1 is lower than L_3 , this means that if star M_2 fills its Roche lobe, matter will pass through L_1 and into the potential well of M_1 . The spatial coordinates are in units of the solar radius, R_\odot .

a white dwarf, neutron star or black hole? Tables 11.1 and 11.2 provide an overview of the range of X-ray binaries that have been discovered. This includes only interacting binaries in which matter

is being transferred from one component to another, thereby making accretion energy available. (Some non-interacting systems in which binarity enhances the normal X-ray output of one

Table 11.1 | Types of X-ray-emitting double stars.

	White dwarf			Neutron star	Black hole
	10 000 km	15 km	5 km		
Magnetic field	Weak	Intermediate	Strong	Weak	Strong
Class	Dwarf novae	Intermediate polars	AM Her systems	Bulge sources (QPOs)	X-ray pulsars
Accretion disc	Yes	Partial	No	Yes	Yes
Companion star	Low mass Cataclysmic variables	Red giant/supergiant Symbiotic stars		Low mass LMXB Bursters	High mass HMXB BeX systems
Typical orbital period	1–10 hours	days–months		11 mins –10 days	A0620-00, Cyg X-1 days –months

Table 11.2 | X-ray binaries.

Type	Donor star	Compact object	Accretion disc	Examples
HMXB	OB I–III	NS, BH	Small	Cen X-3; Cyg X-1
BeX	Be	NS	Small	A0535+26
LMXB	K-M V	NS, BH	Yes	Sco X-1
LMXB	A-F V	NS, BH	Yes	Her X-1; Cyg X-2
LMXB	degenerate	NS	Yes	4U1820-30
CV (dw. nova)	K-M V	WD	Yes	U Gem; SS Cyg
CV (polar)	K-M V	Mag. WD	No	AM Her
CV (IP)	K-M V	Mag. WD	Truncated	DQ Her

Note. NS = neutron star; BH = black hole; WD = white dwarf; CV = cataclysmic variable (see Chap. 10).

member, as in active stellar coronae, have already been described in Chapter 6).

In fact, only one CV, EX Hya, was identified by Uhuru, but because of poor spatial location (and hence a variety of candidate objects), it was not confirmed until much later. Hence all the observational and theoretical work in the Uhuru era concentrated on the neutron star–black hole systems. They represented the standard *X-ray star* and accounted for many of the remarkable properties which attracted attention after the discovery of Sco X-1. The important point is that the enormous gravitational potential of neutron stars and black holes can yield very high luminosities from

relatively small amounts of in-falling material (see Box 11.3). Remember it was the high X-ray luminosities that were completely unexpected.

A typical neutron star system has an object the size of a small city (e.g. Oxford) with a mass equivalent to that of the Sun but radiating in X-rays alone 10 000 times the energy of the Sun at all wavelengths! The amount of matter that needs to fall into the potential well to produce this luminosity is only about $10^{-8} M_{\odot} \text{ yr}^{-1}$, but that is still a staggering amount at a trillion tonnes per second (equivalent to the entire Earth every 300 years)!

In comparison, the white dwarf X-ray binaries are less luminous simply because the compact

Box 11.3 | The power of accretion

Consider a compact object of mass M accreting material at a rate dM/dt ($\equiv \dot{M}$); the luminosity, L , resulting from this accretion will simply be the rate at which gravitational energy is released:

$$L = G M \dot{M} / R, \quad (11.6)$$

where it is assumed that most of this energy is liberated near the object's surface (of radius R). For an M of 1 solar mass, \dot{M} need only be about $10^{-8} M_{\odot} \text{ yr}^{-1}$ to produce a luminosity of $10^{38} \text{ ergs s}^{-1}$, which is close to the maximum seen from any galactic X-ray source.

It is useful to compare the efficiency of this energy-generating process for the three types of compact object with that for nuclear energy. The luminosity obtained from the given \dot{M} may be expressed as a fraction of the total energy the matter possesses (according to Einstein's well-known formula $E = mc^2$), i.e.

$$L = \eta \dot{M} c^2, \quad (11.7)$$

where $\eta = G M / R c^2$; then we find that

- $\eta \sim 0.1$ for neutron stars
- 0.06–0.42 for black holes
- 0.001 for white dwarfs
- 0.01–0.001 for nuclear reactions

which means that accretion onto neutron stars and black holes is the most efficient mechanism known for converting matter into energy. Indeed, we now know that accretion-derived radiation dominates in the observable Universe over all other processes.

object is much less dense. Even though there are many, many more white dwarf systems than classical X-ray binaries, they are only seen at all because of their proximity to us (almost all those known are within a few hundred parsecs, whereas the neutron star–black hole X-ray binaries can now be studied in neighbouring galaxies at distances of millions of parsecs). They are correspondingly much rarer.

11.1.3 Distribution of X-ray sources in the Galaxy

Sky surveys in the early decades of X-ray astronomy quickly revealed that there are only a few hundred luminous X-ray binaries in the Galaxy. They are rare considering that half of all stars in the Galaxy are double stars, and so their formation must require particular fine-tuning. Their distribution on the sky (Fig. 11.6) was also very different for HMXBs and LMXBs. The mass donors in HMXBs have very short lifetimes and so will be found close to their formation site in the spiral arms of the Galaxy. However, the much longer lived LMXBs will be associated with older stellar populations in the central regions of the Galaxy. This is clearer in the ‘face-on’ view of the Galaxy in Fig. 11.7.

11.2 | High-mass X-ray binaries

11.2.1 Discovery of X-ray pulsars

It was the discovery by Uhuru of the eclipsing X-ray source Cen X-3 (Fig. 11.3) which demonstrated beyond any doubt that, for many galactic objects, we are dealing with a binary star phenomenon. The X-ray pulsations with periods of order of seconds also provided valuable clues as to the nature of the compact object. Uhuru found several more eclipsing pulsars, including Vela X-1 and 4U1700-37, and after three more decades of X-ray surveys, there are now around 100 X-ray pulsars known. The great majority of these are associated with massive, early-type stars, usually on the basis of the known orbital period. Details of the observed parameters of some of these X-ray pulsars are contained in Table 11.3.

These systems are of particular importance in X-ray astronomy because they enable the mass of the neutron star to be determined directly by observation. This is possible through the measurement of the velocity of both members of the binary. The spin period of the pulsar exhibits two kinds of modulations to its basic clock or spin rate. The first is a gradual and continuous decrease in the spin period (i.e. the neutron star is being ‘spun-up’; see the next section). The second is a regular variation which is due entirely to the orbital

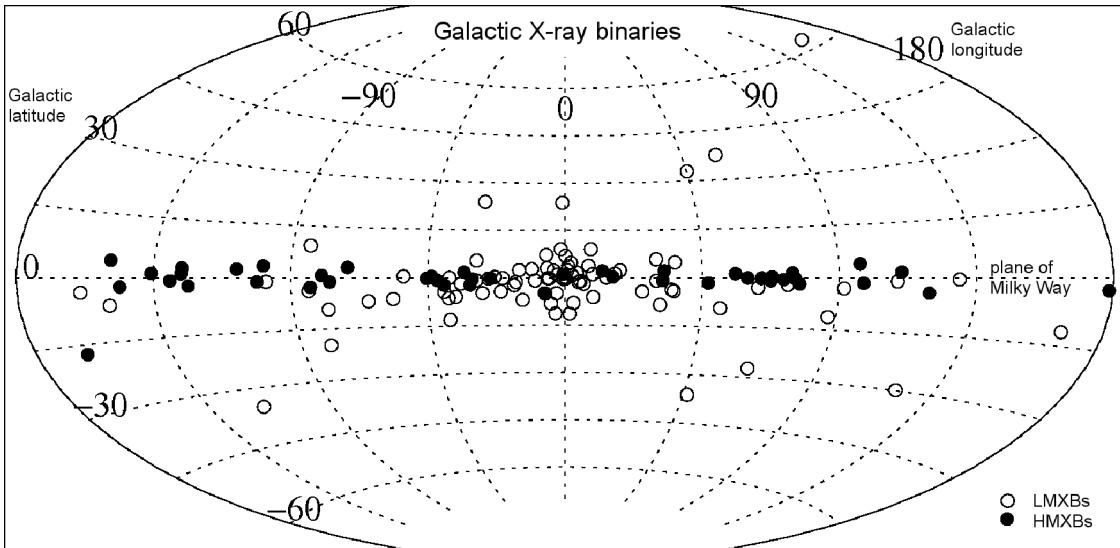
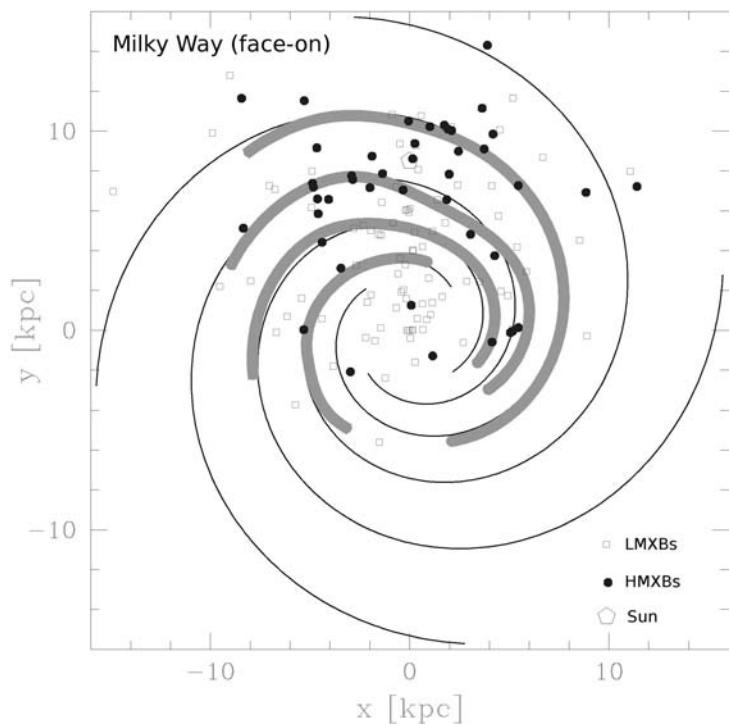


Fig 11.6 Distribution of X-ray binaries on the sky, plotted in galactic coordinates. HMXBs (filled circles) are located in the spiral arms, in the plane of the Galaxy. LMXBs (open circles) are much older and are in the Galactic Bulge (diagram by Grimm *et al.*, 2002).

Fig 11.7 Face-on view of the distribution of X-ray binaries in the Galaxy. Symbols are as in Fig. 11.6, but with the addition of the location of the Sun (the pentagon). The principal spiral arms are shown as gray bands (diagram by Grimm *et al.*, 2002).



motion of the neutron star about its companion (Fig. 11.8).

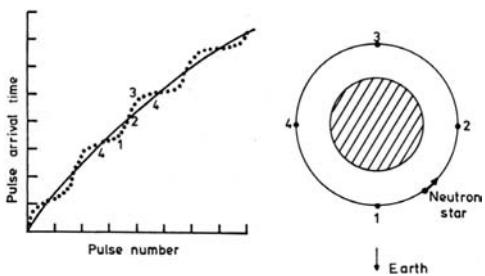
The mass-losing stars in these systems are all optically bright and of early spectral type (usually

earlier than B2), making them easy to observe at high spectral resolution. It is therefore straightforward to measure the Doppler shift of the absorption lines of the primary through the binary orbits.

Table 11.3 X-ray pulsar orbital parameters.

Source ^a	P_{orb} (days)	$a_x \sin i$ (light-second)	Mass function (M_{\odot})	Eccentricity
LMC X-4	1.408	26.0	9.4	-
Her X-1	1.700	13.1831	0.85	<0.0003
Cen X-3	2.087	39.664	15.386	0.0008
4U1538-52	3.728	52.8	11.4	-
SMC X-1	3.892	53.46	10.84	<0.0007
4U1907+09	8.376	83	8.8	0.22
Vela X-1	8.964	113.0	19.29	0.092
4U0115+63	24.309	140.13	5.007	0.3402
2S1553-542	30.6	164	5.0	-
V0332+53	34.25	48	0.101	0.31
GX301-2	41.508	371.2	31.9	0.47

^a see Liu et al. (2006) for a complete catalogue of HMXBs plus detailed references.

**Fig 11.8** Schematic showing the principle of Doppler delay.

A spinning neutron star in a HMXB is shown at right. At left is plotted the time of arrival of each X-ray pulsation against the sequential number of each pulse. The line has a gradual steady increase because the pulsar is spinning up (i.e. rotating faster), but there is a sinusoidal variation (the dotted curve) superposed upon this which reveals the orbital motion. The pulses reach us sooner when the neutron star is at position 1 than when it is at position 3 (diagram by Saul Rappaport and Paul Joss, MIT.)

Such a system is termed a *classical double-lined spectroscopic binary* (see Box 11.4). The one difference with classical astronomy is that one of the components is observed entirely in X-rays and not in the optical. Since they are eclipsing, the inclination of the system is known to a high degree of accuracy, and we can therefore solve directly for all binary parameters (i.e. masses of both components, their separation and the physical size of the mass donor).

There are a number of X-ray binary systems for which such a solution is possible, and the results for six of the most well known are given in Table 11.4, with Fig. 11.10 showing the quality of the data available for SMC X-1 and 4U0115+63.

The sizes of the orbits and companion stars are shown to scale in Fig. 11.11, and their measured masses are compared with (the much more

Box 11.4 | Measuring neutron star masses

Determining stellar masses in binaries by measuring their relative velocities and the size of the orbit has been the fundamental method by which astronomers have determined stellar masses for almost 100 years. (Indeed, masses for stars *not* in binary systems can *only* be estimated by indirect means and are notoriously inaccurate.) Radial velocity studies can easily be applied to X-ray

binaries in which the mass-losing star is observable in the optical and the compact object is an X-ray pulsar. The velocity of each can be determined by the Doppler effect, one in the optical, the other in X-rays (since the observed pulsar period will change smoothly through the orbit by the same process, as demonstrated in Fig. 11.3).

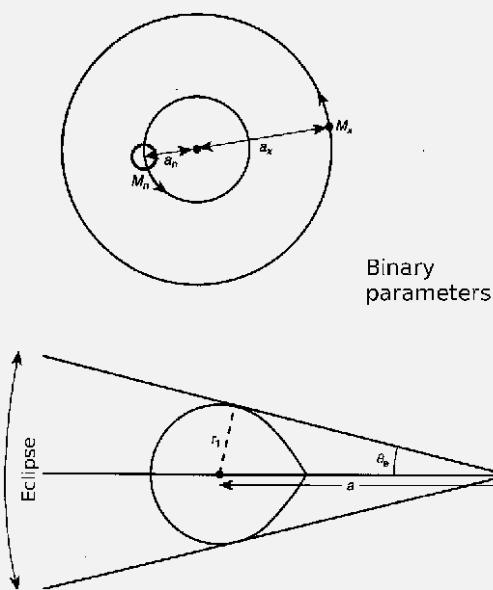


Fig 11.9 (top) Definition of parameters in a binary system and (bottom) the geometry of the eclipse of the X-ray source by its companion star.

Referring to Fig. 11.9 and assuming for simplicity a circular orbit, application of Kepler's third law gives

$$(a_x + a_n)^3 / P^2 = G(M_x + M_n) / 4\pi^2, \quad (11.8)$$

and the centre of mass is calculated from $M_x a_x = M_n a_n$. But we actually observe $a_x \sin i$ (i is the inclination angle between the orbital plane and the plane of the sky) and $a_n \sin i$ through measuring

the so-called K velocities of the optical and X-ray stars:

$$K_x = 2\pi a_x \sin i / P \quad (11.9)$$

(since $v = 2\pi r / P$) and

$$K_n = 2\pi a_n \sin i / P. \quad (11.10)$$

It is therefore immediately possible to calculate the mass ratio of the two stars, $q (= M_x / M_n)$, but without some indication of the value of i we cannot find the individual masses. This is demonstrated by introducing the mass function, $f(M)$, defined from the preceding equations as

$$f_x(M) = \frac{M_n^3 \sin^3 i}{(M_n + M_x)^2} = \frac{4\pi^2 a_x^3 \sin^3 i}{G P^2} = \frac{P K_x^3}{2\pi G}, \quad (11.11)$$

which is obtained entirely from the observable quantities P and K_x but which needs i to solve for the masses themselves. An estimate for i is forthcoming in the eclipsing systems from the value of the eclipse half-angle, θ_e . Using the preceding geometry, we obtain

$$r_1/a = (\cos^2 i + \sin^2 i \sin^2 \theta_e)^{1/2}, \quad (11.12)$$

where a is $a_x + a_n$, the total separation of the stars. The average size of the optical star, r_1 , can be estimated from the spectral type, and r_1/a has been computed as a function of q (see Box 11.2). Hence for a handful of systems it is possible to obtain M_x and M_n directly, and these are the values given in Table 11.4.

Table 11.4 HMXB neutron star masses.

Source	Companion mass (M_\odot)	Companion radius (R_\odot)	i (degrees) (M_\odot)	Neutron star mass
Her X-1	1.99	3.86	80	0.98
SMC X-1	16.8	16.3	65	1.06
Cen X-3	19.8	12.2	75	1.06
LMC X-4	14.7	7.57	68	1.38
Vela X-1	23.0	34.0	83	1.77
4U1538-52	16.9	15.2	71	1.8

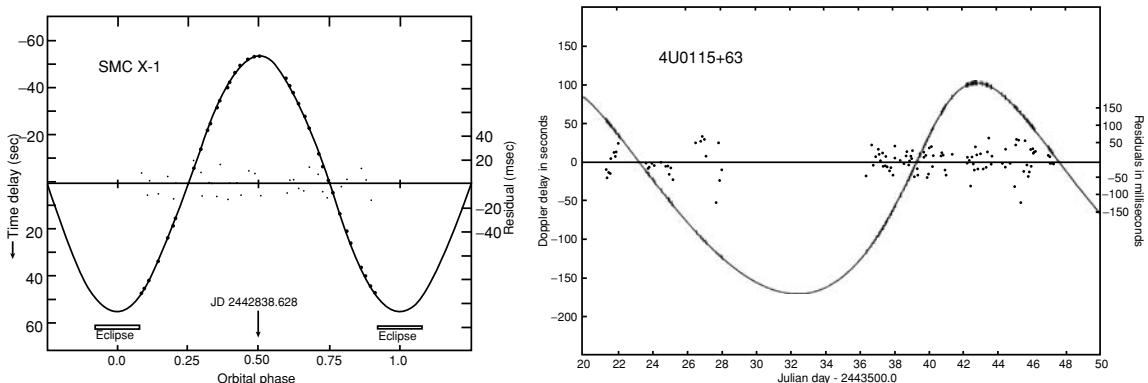


Fig 11.10 | Orbits of two massive X-ray binaries, SMC X-1 and 4U0115+63, as determined from observations of the Doppler delay. SMC X-1 is essentially circular, but 4U0115+63 is notably elliptical. The quality of the data is evident from the residuals (i.e. difference between the observations and orbital fits), which are displayed at a scale 1000 times greater than the observations themselves (diagram by Saul Rappaport and Paul Joss, MIT).

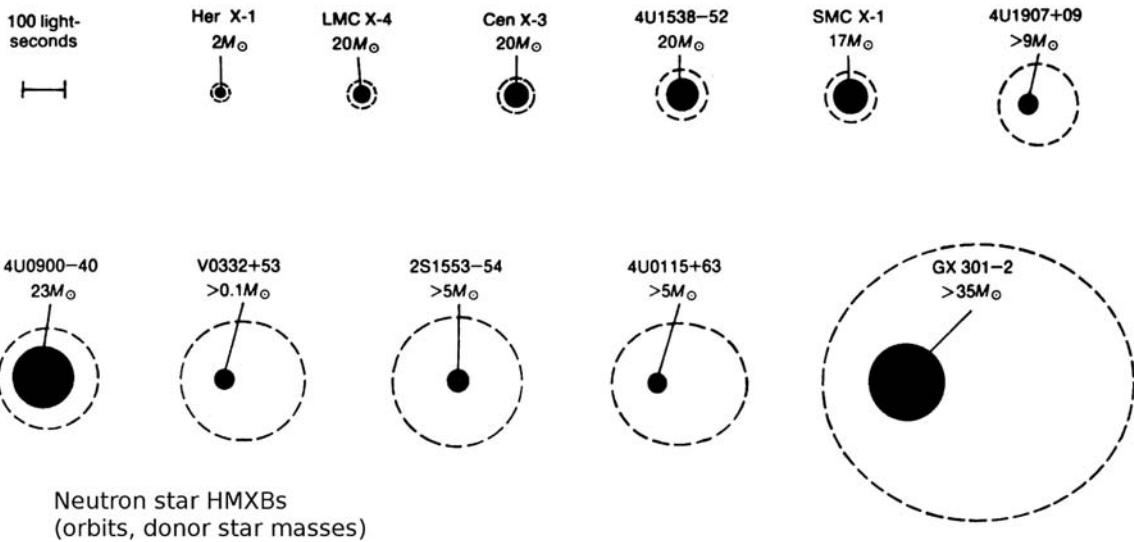
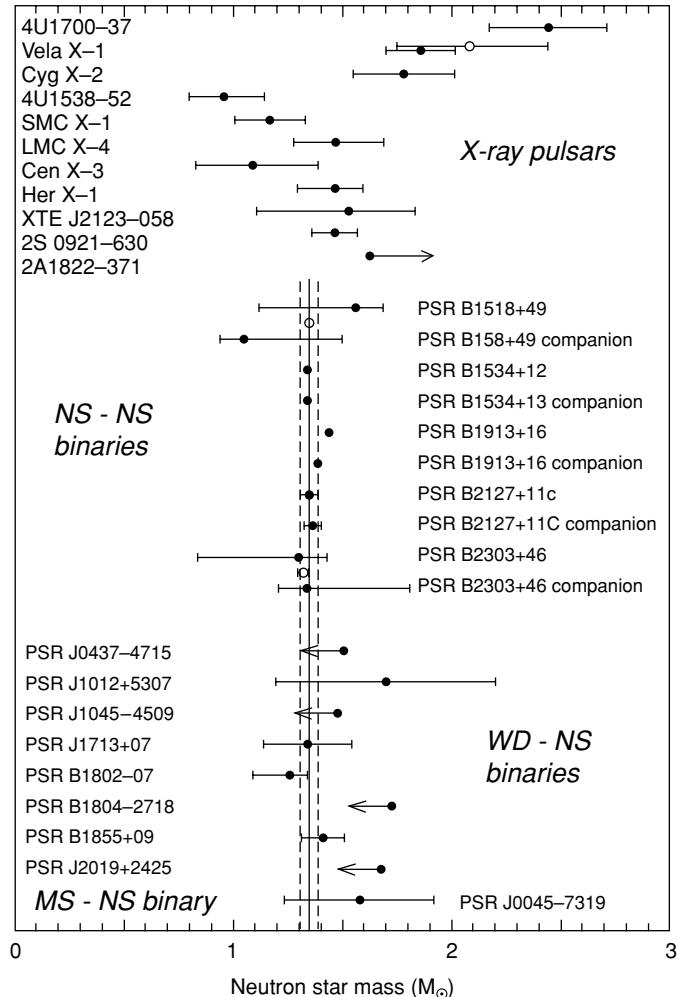


Fig 11.11 | Orbits to scale for a selection of the massive X-ray binaries, as determined from optical and X-ray observations. Note that the eccentric binaries are both larger and have the longest periods (10–40 days). The bar denotes the size of 100 light-seconds (diagrams based on originals by Saul Rappaport and Paul Joss and modified by Los Alamos National Laboratory (LANL) and F. Nagase).

accurate) radio pulsar masses in Fig. 11.12. The range of uncertainty for these mass calculations is indicated by the error bars in which all the measurement errors have been taken into account. Note that the component masses of PSR1913+16, the first and most extensively studied binary radio pulsar, still represent the most accurate neutron star masses known and define the canonical neutron star mass of $1.4 M_{\odot}$.

Accurate mass determinations have an impact way beyond the study of X-ray pulsars. The maximum allowed mass of a neutron star is an important parameter that tells us the equation of state of the material inside the neutron star (i.e. how its pressure is related to the density and temperature) and, because of the very high densities involved, brings us directly into contact with fundamental particle physics. One of the permissible equations

Fig 11.12 X-ray pulsar masses (upper section) are here compared with neutron star masses as determined from radio pulsar observations of double neutron star binaries, white dwarf – neutron star binaries and a main sequence star – neutron star binary (lower sections). It is remarkable how close the measured neutron star masses are to the canonical value of $1.4 M_{\odot}$, although there is now a hint that this may be exceeded by at least some of the accreting systems (diagram compiled and updated from Thorsett & Chakrabarty, 1999; Lattimer & Prakash, 2005).



of state predicts a maximum neutron star mass of just over $1.5 M_{\odot}$ and may already be excluded by the estimated mass of several X-ray pulsars. Clearly even more accurate determinations of these and other neutron star binaries are necessary. These results will also be of considerable importance in comparison with putative black-hole masses in Chapter 12.

11.2.2 Spinning X-ray pulsars up and down

Accurate masses, based on purely dynamical considerations, have now been measured for these systems, but what is the actual accretion process at work? The basic model was developed by Davidson and Ostriker (1973) shortly after the first binary

signature of an HMXB was established. It is shown schematically in Fig. 11.13 (see also Fig. 11.4).

UV observations of early-type stars show that such stars have prodigious stellar winds. Values as high as 10^{-5} to $10^{-4} M_{\odot} \text{ yr}^{-1}$ are known, but 10^{-6} is typical. The wind is driven by the extremely high luminosity of these massive, early-type stars (they can be 10 000 times brighter optically than the Sun; for comparison, the Sun's wind carries away a mere $3 \times 10^{-14} M_{\odot} \text{ yr}^{-1}$). Because of the high temperatures of early-type stars, they are powerful UV emitters, and it is the UV photons that are absorbed and scattered readily by surrounding material, thereby imparting momentum and driving the expanding wind. This high mass loss occurs in all directions from the surface of the star, and the orbiting neutron star has to pass

Box 11.5 | Bondi-Hoyle accretion

Bondi-Hoyle accretion is the calculation of the amount of material accreted by a body travelling through a uniform density medium such as interstellar space (Bondi & Hoyle, 1944). It can be applied quite straightforwardly to a compact object moving in a stellar wind, as is the case for the massive X-ray binaries. The basics of the geometry of the problem are shown in Fig. 11.13.

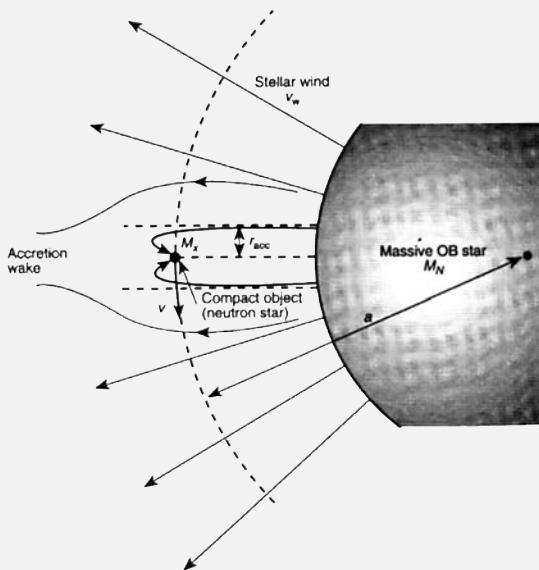


Fig 11.13 Geometry of accretion onto a compact object orbiting its massive companion. The wind of material leaving uniformly in all directions from this star's surface will accrete onto the neutron star only if it is within a critical distance (r_{acc}) of the compact object, thereby forming an *accretion cylinder*. Material that is perturbed in its flow but does not actually accrete forms an *accretion wake* beyond the compact object.

It is assumed that the massive star possesses a wind that flows out uniformly in all directions. Material passing within a distance r_{acc} from the

through it. The calculation of how much of this material will be captured by the compact object was first performed by Bondi and Hoyle (in 1944!) and has been applied in this case (see Box 11.5). Even though only about 0.1 per cent of the wind is captured by the orbiting star, this is sufficient

compact object will be accreted, whereas material outside this cylinder will escape. This radius is calculated by noting that material will only be accreted if it has a kinetic energy less than the potential energy in the vicinity of the compact object (of mass M_X), i.e. it is set by

$$\frac{1}{2}mv_{\text{rel}}^2 = \frac{GM_Xm}{r_{\text{acc}}} \quad (11.13)$$

for a particle of mass m , which gives

$$r_{\text{acc}} = \frac{2GM_X}{v_{\text{rel}}^2}, \quad (11.14)$$

where v_{rel} is the relative velocity of the compact object and the stellar wind. This is computed from

$$v_{\text{rel}}^2 = v^2 + v_w^2, \quad (11.15)$$

where $v^2 = GM_N/a$. The normal (early-type) star has mass M_N , and a is the radius of the orbit. The amount of material accreted by the compact object is then given by the amount inside the accretion cylinder, which is

$$\dot{M} = \pi r_{\text{acc}}^2 v_{\text{rel}} \rho, \quad (11.16)$$

where the density ρ can be calculated from the assumption that the wind (stellar mass-loss rate \dot{M}_w) is uniform, i.e.

$$\rho = \frac{\dot{M}_w}{4\pi a^2 v_w}. \quad (11.17)$$

These equations can be solved to give the fraction of the wind that is accreted onto the compact object:

$$\frac{\dot{M}}{\dot{M}_w} = \left(\frac{M_X}{M_N} \right)^2 \frac{(v/v_w)^4}{[1 + (v/v_w)^2]^{3/2}}, \quad (11.18)$$

which yields values of order 10^{-3} to 10^{-5} for typical HMXBs.

to power a significant fraction of the observed X-ray luminosity. However, there are cases, including Cen X-3, in which the observed wind mass-loss rate is insufficient to power the source entirely, and an additional mechanism is needed to give the neutron star the fuel it needs. This almost

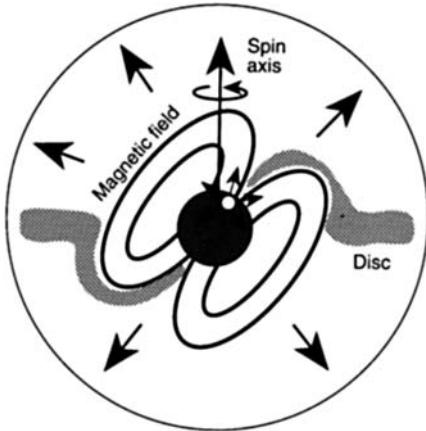


Fig 11.14 Schematic of an X-ray pulsar and its magnetosphere. At the magnetospheric boundary, the in-falling material from the accretion disc is forced to travel along the magnetic field lines to the polar caps of the spinning neutron star. It is near the polar caps that the X-rays are generated. If, as here, the spin and magnetic axes of the neutron star are not aligned, then the rotation of the neutron star will naturally give rise to the periodic modulation that we see as an X-ray pulsar.

certainly arises from the evolution of the early-type star towards filling its Roche lobe, whereby matter spills over the inner Lagrangian point directly onto the compact object (as in LMXBs).

The next step is to ask how the accreted gas finds its way onto the neutron star surface and what is the physical origin of the pulsations? Within the capture radius of Fig. 11.13, the material will probably form an accretion disc, but this will be limited as a result of the material's small angular momentum. However, as the gas moves through the disc towards the neutron star, it will reach a point at which its motion is controlled by the powerful magnetic field of the collapsed star. This point is called the magnetospheric radius, as shown in Fig. 11.14.

The gas will now move along the field lines towards the magnetic polar caps of the neutron star. As material hits the neutron star surface, a very hot shock is formed in which X-rays are produced. Because of the existence of the column of material above it, the X-ray emission is not uniform in all directions but shadowed into a kind of fan beam. To produce a modulation of the X-rays at the neutron star spin period, it is necessary for

these magnetic polar caps to be displaced from the rotation axis. In this way a lighthouse effect occurs and regularly changes the angle of view of the X-ray region.

These X-ray pulsars tend (on average) to be spinning up (Fig. 11.15) because the accreted material from the companion star has angular momentum which is ultimately transferred to the neutron star. A simple extension of the Bondi-Hoyle accretion model can be used to calculate the magnitude of the expected spin-up rate, and this is compared with the observed rates for two well-studied sources in Fig. 11.16. The key point is that the rate increases with the luminosity of the source, as expected given that the luminosity is proportional to the accretion rate, and more material will carry more angular momentum.

Many of these sources, however, have periods of inactivity when they are weak or undetectable in X-rays. We presume that the accretion rate is much lower during these periods and so spin-up will not occur. Indeed, we know from observations of single radio pulsars that they spin down with time as a result of the energy loss in accelerating relativistic particles (see Chapter 9). It is therefore to be expected that both spin-up and spin-down will be seen to occur, although many of the individual observations are peculiar and not well understood.

11.2.3 Accreting from a stellar wind

A luminous X-ray source close to its companion will have a significant effect on the behaviour of the star's atmosphere and wind. The source itself can be used to probe the structure of the wind by observing how the X-ray flux and spectrum changes as the compact object is eclipsed by its supergiant companion. The best examples of this are 4U1700-37 and Cen X-3, which also cover a large range of luminosity, Cen X-3 being almost 100 times brighter than 4U1700-37. The X-ray source is seen through different parts of the primary's wind and atmosphere as the eclipse progresses (Fig. 11.17).

If the wind is dense and not highly ionised, then it can obscure low-energy X-rays, thereby changing the observed spectrum of the source. We characterise this as a change in the column

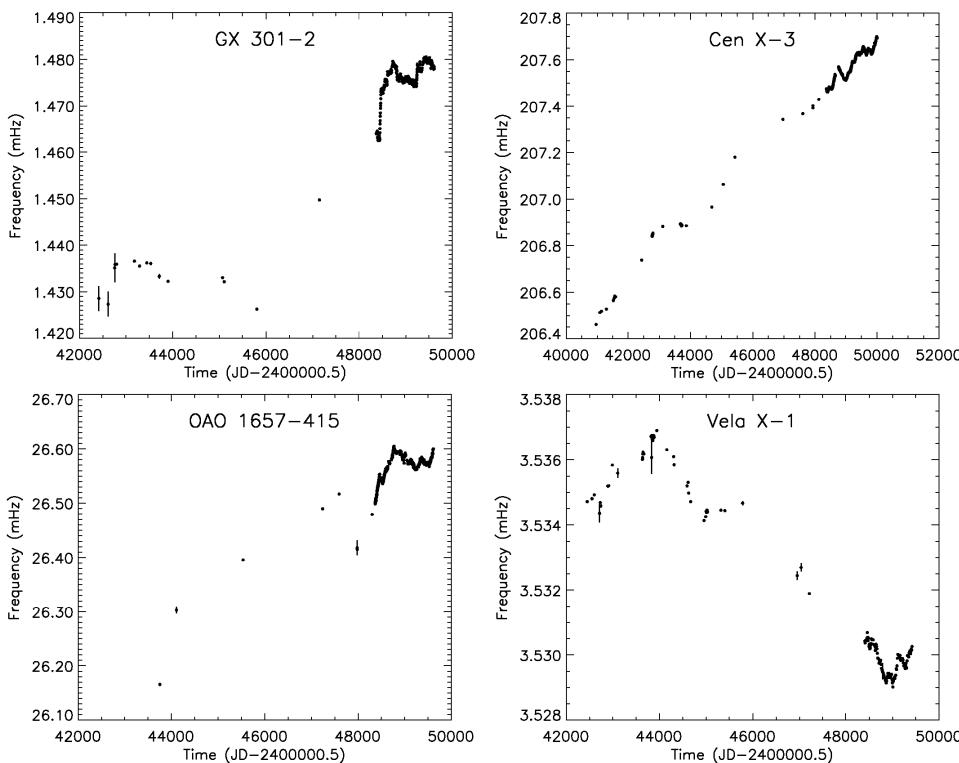


Fig 11.15 The change in period of a selection of X-ray pulsars is shown here over timescales of decades, from Uhuru in the 1970s to the Burst and Transient Source Experiment (BATSE) in the 1990s. While the general trend is one of spinning up (i.e. increasing rotation frequency), there are substantial deviations (sometimes quite wild) from this trend (e.g. Vela X-1) (diagram from Bildsten *et al.*, 1997).

density of material between us and the X-ray source. On the other hand, if the X-ray source is extremely bright, it is capable of completely ionising all the material in the wind. Such ionised gas cannot absorb X-rays, and the spectrum will be little affected as eclipse approaches, with the source brightness dropping abruptly as it is finally occulted by the primary. The long (days) binary periods of HMXBs mean that long observing runs are required to study these effects, preferably with no break in coverage near the eclipse itself. EXOSAT was an ideal mission to undertake such observations, and Fig. 11.18 shows the light curve of 4U1700-37.

Outside eclipse, 4U1700-37 shows chaotic variability, yet the average luminosity of this object is consistent with that expected for accretion from a wind (see Box 11.5). Hence the variability is related to the structure or inhomogeneity present

in the wind. The X-ray spectra enable the column density, N_X , to be calculated, and this reveals one of the gross features of accretion from a stellar wind. This is the increase in N_X that occurs just after phase 0.6, shortly after the neutron star has passed in front of the primary (remember that eclipse of the X-rays occurs from phase 0.9 to 0.1). It is best explained as being part of an additional gas stream from the primary that misses the neutron star (see Fig. 11.19), leaving a trailing ‘wake’. The reason for this is that the primary is not co-rotating with the neutron star. Normal tidal lobe overflow (transferring material directly from one star to the other, as in the LMXBs) will not occur.

11.2.4 Be stars and X-ray pulsars

It was noted by Maraschi *et al.* (1976) that HMXBs also seem to divide into two groups: those with

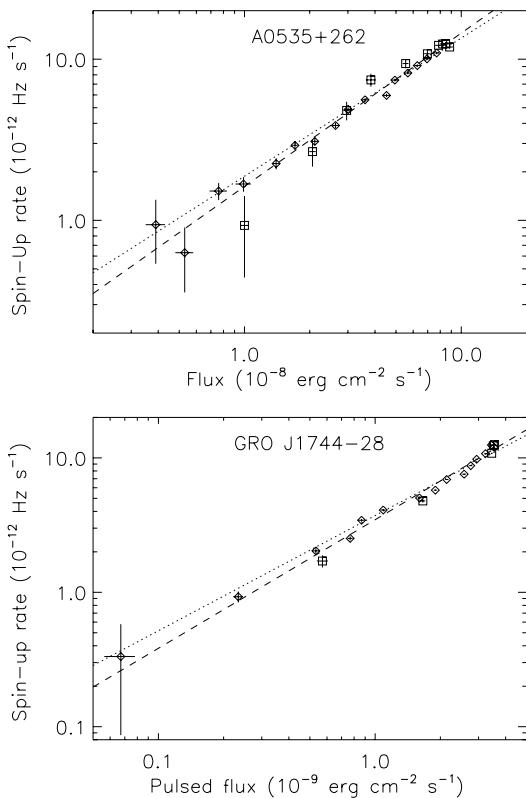


Fig 11.16 The spin-up rates (i.e. rate of increase in the spin periods) of two X-ray pulsars monitored by BATSE, as a function of the observed X-ray flux. Crudely put, the higher the X-ray luminosity, the more material is accreted by the pulsar and hence the faster it is spun up. The harder you hit a spinning top, the faster it will spin (diagram from Bildsten et al., 1997)!

OB supergiant primaries showing characteristics of powerful stellar wind mass loss (such as Cen X-3) and those with mid-B giant or main sequence primaries that were classified as Be stars. Now Be stars are a famous class of stars that have been known since the early twentieth century and were studied intensively by Otto Struve in the 1930s. The e stands for ‘emission lines’, which are usually variable in both intensity and profile on short and long timescales (hours to years). Their nature has been a subject of speculation for two generations of astronomers. The fact that some of them are in binary systems with neutron star companions (pulsars) has greatly renewed interest in this class of stars.

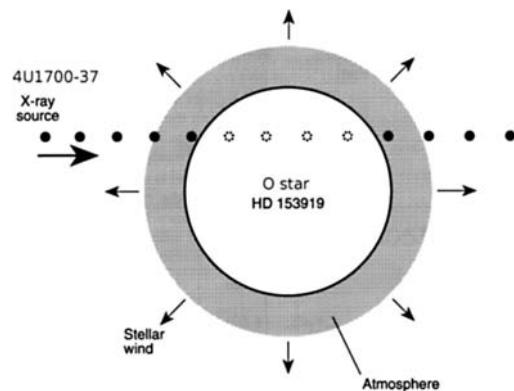


Fig 11.17 Schematic of the eclipse of 4U1700-37 by its massive O star companion, HD153919. The eclipse is not sharp because the hot star is driving a powerful wind of outflowing material. This material becomes more opaque to X-rays the closer it is to the stellar surface. It is possible then to estimate the density and structure of the atmosphere by observing the change in X-ray intensity and spectrum as the source goes into and then comes out of eclipse. Such observations (of 4U1700-37) were undertaken by EXOSAT, and the light curve is shown in Fig 11.18.

With the long-term X-ray surveys by various satellites over the last 30 years came the discovery that major outbursts of some of the Be systems were in fact repeating, but with very long periods. Details of this group are summarised in Table 11.5 using some well-known examples.

The Be phenomenon is almost certainly due to episodes of mass expulsion from the star, probably around the equator, so as to form an equatorial ring or disc. This is likely associated with Be stars being rapid rotators and hence rotating close to their breakup speed, making it easier to form the surrounding ring. Any compact object orbiting such a star would encounter the ring, thereby producing a large increase in its X-ray output (Figs. 11.20 and 11.21). If, in addition, the orbit is eccentric, then the X-ray output will be modulated on the orbital period, and by a much larger factor than could be accounted for by a simple stellar wind from the B star. Indeed, in almost all cases, the wind from a B star could not possibly give the observed X-ray luminosities from these systems.

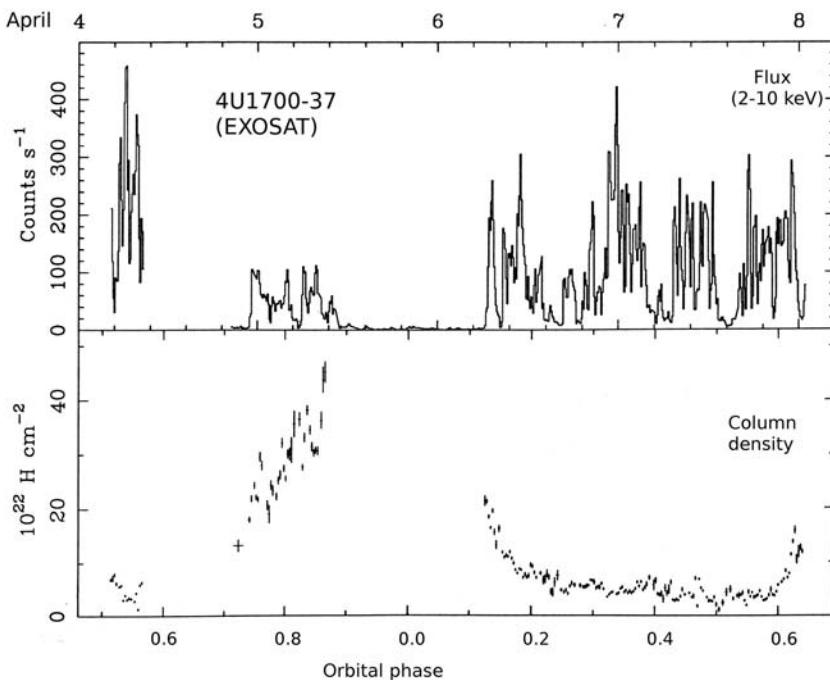


Fig 11.18 An unbroken 80-hour observation of 4U1700-37 by EXOSAT in April 1985, with (top) X-ray flux and (bottom) absorbing column density as inferred from the X-ray spectrum. The X-ray-emitting neutron star is eclipsed by its supergiant companion from phases 0.9–1.1 but shows highly chaotic variability outside eclipse. This is almost certainly due to substantial structure (clumpiness) in the wind emanating from the supergiant. Note the asymmetry in the absorption prior to eclipse (phases 0.7–0.9) (courtesy of EXOSAT Observatory).

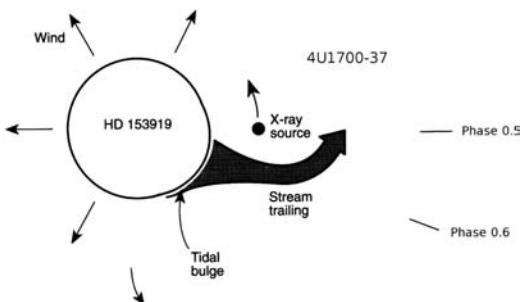


Fig 11.19 Schematic of the enhanced absorption at phase 0.6 in 4U1700-37 (Fig. 11.18). The massive star and X-ray source are not co-rotating (i.e. it does not rotate once for every orbit) and so a tidal bulge could give rise to a stream that misses the neutron star and trails behind it, thereby giving rise to the extra absorption.

There is strong support for this interpretation from the distribution of X-ray pulsar periods as a function of their binary periods (see Fig. 11.22). The supergiant systems

yield a relationship which is almost constant with orbital period and consistent with that expected from the known stellar wind variation of density. The Be systems, however, show a strong correlation (Corbet, 1984) that is very different from that of a stellar wind. Basically, they are consistent with a much slower wind, much as would be expected from an equatorial ring around a Be star. The larger orbital periods require a greater separation of the neutron star and Be companion, which implies less matter to spin up the neutron star and hence a longer spin period, as observed.

Of course, because of the episodic nature of the mass ejection, the pulsar spin period will not stay constant. During outbursts it will speed up, and during periods of quiescence it will gradually spin down. But eventually, each system will reach a balance between the two, and that is assumed to be true for these objects now as they show very little change during outburst.

Table 11.5 | X-ray emitting Be systems.

Source	Spectral type	V (mag)	P_{spin} (s)	P_{orb} (days)	e	$L_X(\text{peak})$ (ergs s^{-1})
A0538-66	B2IIIe	15	0.069	16.65	0.7	10^{39}
4U0115+63	Be	16	3.6	24.3	0.34	8×10^{36}
V0332+53	O8.5V	15.5	4.4	34	0.31	1.5×10^{38}
EXO2030+375	B0Ve	19.7	41.8	46	0.41	10^{37}
A0535+26	B0Ve	9	104	111	0.3	2×10^{37}
GX304-1	B1Ve	14	272	133	>0.5	3×10^{36}
4U1154-61	B1Ve	9	292	188	>0.5	6×10^{36}
X Per	O9.5III-Ve	6	835	250	0.11	10^{35}

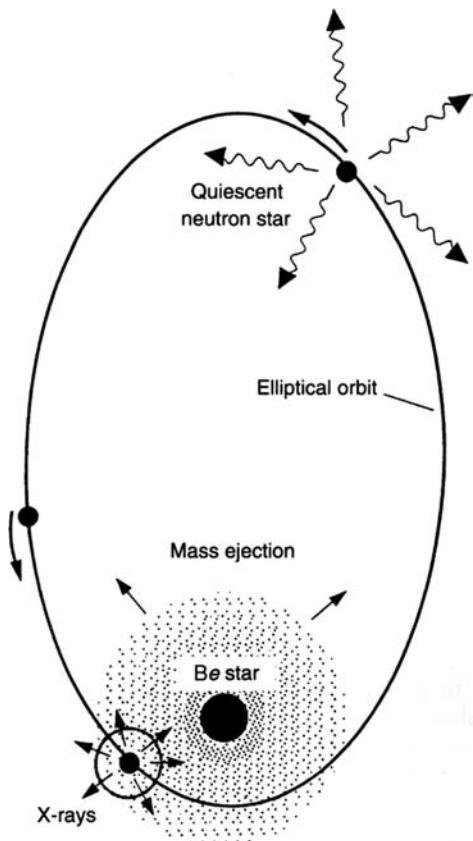


Fig 11.20 Schematic of a Be X-ray binary consisting of a Be star whose rapid rotation gives rise to an extended surrounding disc or ring of material. If the orbiting neutron star has a high eccentricity, this means it will only encounter this ring during periastron passage, at which times the accretion of material leads to an X-ray transient outburst. The enormous X-ray luminosity released during these passages can drive much of this extended ring material out of the system (diagram by Ed van den Heuvel).

11.2.5 X-ray spectroscopy of HMXBs

The substantial hard X-ray emission produced by the accreting neutron star will have a profound influence on the outflowing stellar wind in which it is embedded (Fig. 11.23). The X-rays will be scattered, absorbed and reprocessed into an emission line and continuum spectrum that contains information about the material in which this has taken place. The X-ray effect on its surroundings is even more profound given that it can destroy (by ionising the wind material) the mechanism by which the wind is physically driven to large outflow velocities (intense UV radiation from the donor, absorbed by matter in the wind). All these effects are very sensitive to the X-ray luminosity level.

The first high-resolution X-ray spectra of HMXBs were obtained by ASCA, from which Liedahl and Paerels (1996) and Kawashima and Kitamoto (1996) showed that, as suspected, the X-ray emission lines seen from these objects were due to photoionisation of the wind material by the X-ray source. Sako *et al.* (1999) went further in their analysis of Vela X-1 to show that the X-ray spectrum required a combination of highly ionised and cold (essentially neutral) matter to explain the data. This meant that the wind must be *inhomogeneous* (i.e. clumpy!). They estimated that these cold clouds contained >90 per cent of the mass of the outflowing wind, but the highly ionised component occupied >95 per cent of the volume.

The broadband X-ray spectral features of HMXBs are nicely summarised in Fig. 11.24. Both Cen X-3 (out-of-eclipse here) and Cyg X-3 have high L_X , but the low-column density to Cen X-3 means

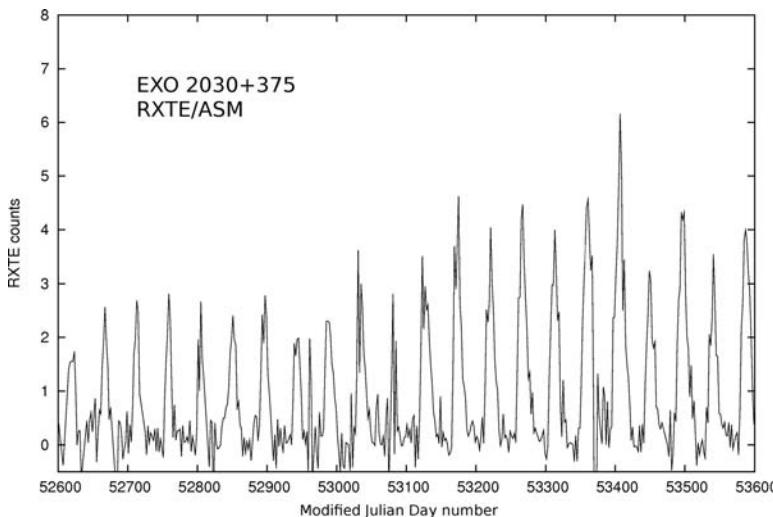


Fig 11.21 RXTE ASM light curve covering more than 2 years of the BeX system EXO2030+375. The X-ray outbursts are periodic (46 days) as the neutron star encounters the equatorial ring of the Be star (adapted from an original by Ignacio Negueruela).

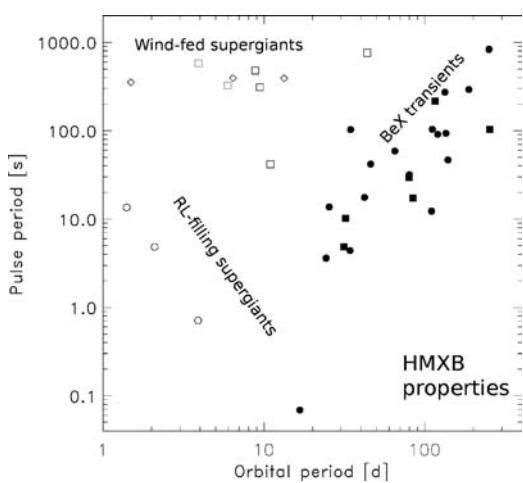


Fig 11.22 A plot of the X-ray pulsar periods of HMXBs against their orbital periods, which divides into three main regions. Upper left (open squares) are supergiant X-ray binaries in which mass transfer is via a stellar wind; lower left (open circles) are supergiants which are filling their Roche lobes; middle right (filled circles) are the Be X-ray transients. Open diamonds and filled squares are X-ray sources that are not yet optically identified (diagram by Robin Corbet).

that the continuum is dominant and the emission lines are weak, whereas the higher density around Cyg X-3 produces much stronger emission lines. In Vela X-1 the low-energy X-ray (longer wavelength) continuum is much more highly absorbed, but the soft X-ray lines are stronger, implying that

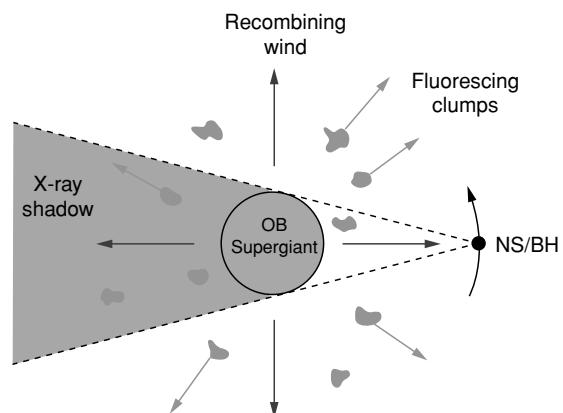


Fig 11.23 Geometry of an HMXB demonstrating the eclipse 'shadow' cone cast by the massive donor star from the X-ray luminous neutron star. The wind consists of a diffuse, low-density outflow in which are embedded clumps of much higher density (diagram from Sako et al., 2003).

they are not affected by the same absorbing material. However, in GX301-2, both soft X-ray emission lines and continua are highly absorbed, leaving the high-energy iron emission as the dominant feature.

Observations of Vela X-1 through different phases of its orbit are shown in Fig. 11.25. While the continuum changes by an order of magnitude, the emission line components (extracted separately in Colour Plate 29) are virtually unchanged

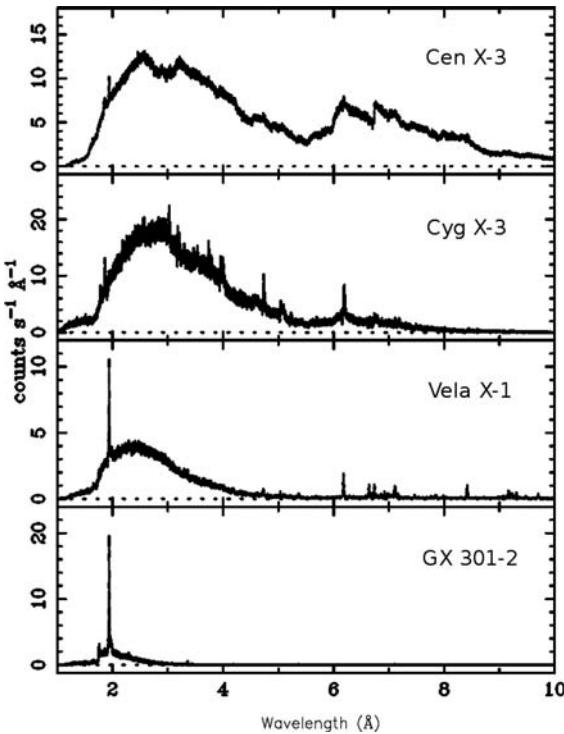


Fig 11.24 Chandra HETGS spectra of four X-ray-luminous HMXBs. Note their very different continuum and emission line characteristics (diagram by Sako *et al.*, 2003).

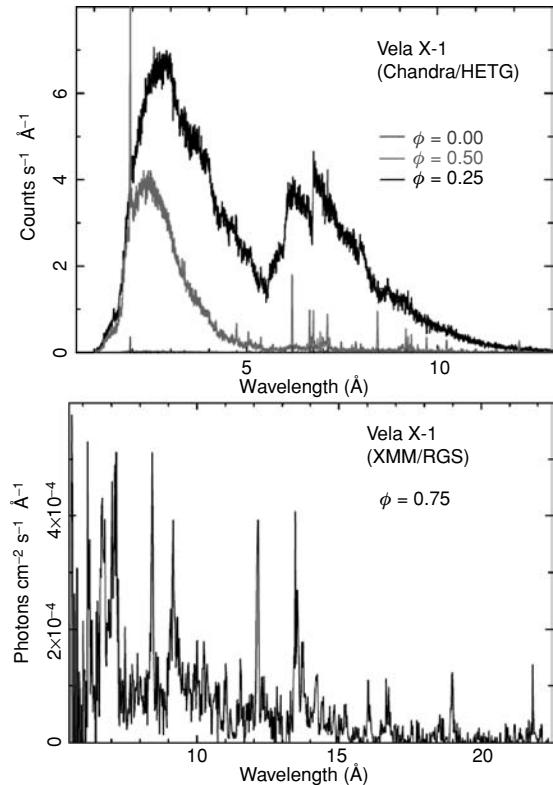


Fig 11.25 Chandra/HETGS and XMM/RGS spectra of Vela X-1 at a variety of phases through its binary cycle (diagram by Sako *et al.*, 2003).

at these opposite phases of the orbit. This is exactly what would be expected if the X-ray source only disrupted the stellar wind close to the compact object.

The diagnostic power of Chandra's high spectral resolution for probing HMXBs is superbly demonstrated in the spectrum (Fig. 11.26) of GX301-2 immediately prior to periastron passage. The spectrum is dominated by the iron K α fluorescence feature at 6.4 keV, but the blowup of the line profile shows that it has a shoulder to lower energies (longer wavelengths). This is a result of Compton scattering of the emission line photons off the electrons in the stellar wind, which reduces the energy of the photon by an amount that depends on the scattering angle (and which is why it produces a shoulder on the line profile). From this profile, it was shown that the electron temperature was low (<6 eV) and consistent with that expected for an expanding stellar wind.

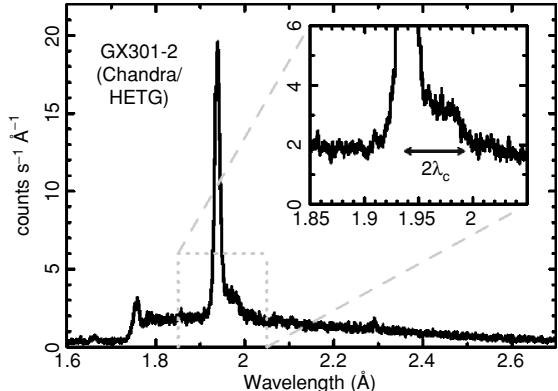


Fig 11.26 Chandra/HETGS spectrum of GX301-2 just prior to periastron passage. Dominated by the FeK α fluorescent line at 6.4 keV, the inset shows that the line profile exhibits a clear shoulder to lower energies. This is attributed to Compton scattering of these X-ray photons off electrons in the immediate environs of the X-ray source (diagram by Sako *et al.*, 2003).

Table 11.6 General properties of HMXBs.

Type ^a	Percentage of all HMXB	Optical luminosity class	Typical pulse period (s)	Typical binary period (days)	Typical binary eccentricity	Log L_X (ergs s ⁻¹)
Be	89	III-V	0.05–500	2–260	0.3–0.9	36–38
SG	11	I-II	200–700	3–40		34–36

^a SG = supergiant systems.

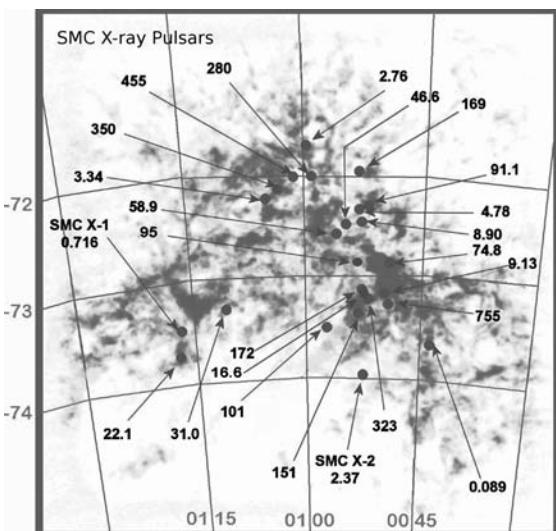


Fig 11.27 Radio (21 cm) image of the SMC, on which are superposed the locations of 25 HMXB X-ray pulsars. The numbers indicate the pulse period in seconds (diagram by Stanimirovic *et al.*, 1999).

11.2.6 Populations of HMXBs

The relative fractions of the different types of HMXB indicated in Fig. 11.22 are given in Table 11.6, which is based on HMXBs observed in our Galaxy and the Magellanic Clouds, particularly the SMC. X-ray surveys of the latter have revealed a very surprising result. The SMC is a small, satellite galaxy of the Milky Way and only contains approx. 2 per cent of the mass of the Milky Way. Consequently, with around 65 HMXBs known in our Galaxy, one would naively expect to find only one or two HMXBs in the SMC. Yet >50 are now known (Fig. 11.27)!

The explanation for this remarkable overabundance of HMXBs in the SMC is linked to new

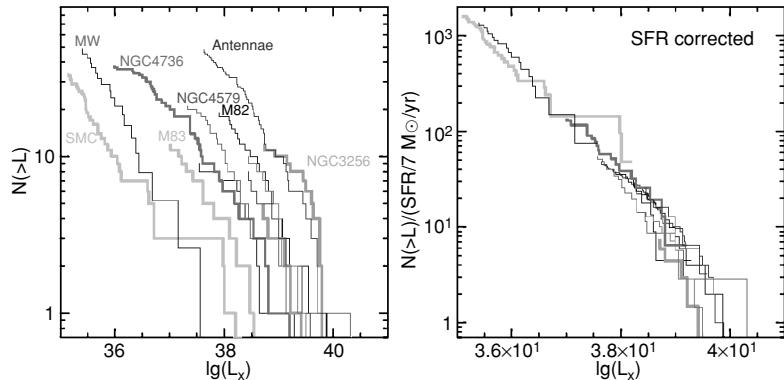
evidence concerning its recent star-formation history. Radio observations have revealed ‘bridges’ of material between the Small and Large Clouds, which were produced by tidal interactions between them during a close approach about 100 Myr ago. The tidal forces triggered a burst of star formation, which we see today as an enhanced number of HMXBs.

The link between HMXBs and star-formation rates in galaxies was made very clear by Grimm *et al.* (2003) when they examined the *luminosity function* of HMXBs in a number of galaxies. This has been made possible by Chandra’s exquisite spatial resolution, which allows the association of X-ray sources with spiral arms and massive star-forming regions to be clearly delineated. The luminosity function is simply the number of X-ray sources brighter than a given luminosity, L , and clearly declines as L increases. The results are plotted in Fig. 11.28 (left) and show that while the galaxies all follow the same basic curve, they are widely separated from each other. However, the remarkable insight brought by Grimm *et al.* was to note that the range of star-formation rates in these galaxies is quite large (from ~ 0.15 to $7 M_\odot \text{yr}^{-1}$) and that each can be brought into almost perfect alignment by scaling each relative to its star-formation rate.

11.2.7 Evolution of an HMXB

It is a well-known phenomenon in stellar evolution that the more massive a star the faster it consumes its fuel and the shorter is its main sequence lifetime. This goes against one’s intuition that a star which has more fuel will last longer. In fact a $10 M_\odot$ star will last only roughly 1/100 as long as our Sun! However, in HMXBs, we find a neutron star (virtually the end point of a star’s life)

Fig 11.28 (left) The high spatial resolution of Chandra allows the luminosity function $N(> L)$ of HMXBs (i.e. simply the number of sources whose X-ray luminosity is greater than L) to be directly measured in a number of galaxies. (right) These same luminosity functions have here been scaled by a factor which is the ratio of each galaxy's star-formation rate to that of the Antennae (diagram by Grimm et al., 2003).



orbiting a much more massive companion star. How can this be? Surely the more massive star would be expected to have evolved first?

The answer to this paradox is that the neutron star in an HMXB actually started out as the *more* massive of the pair, but it transferred the greater part of its mass to its companion towards the end of its very short lifetime, as it began its evolution into what would have been (for a single star) the giant phase. Computer models allow the evolution of such a system to be followed, and an example is given in Fig. 11.29.

The simulation starts at time 0 with a massive pair of stars in a 100-day binary. The heavier star evolves rapidly through normal hydrogen burning, generates a helium core and, after 13 million years, expands to fill its Roche lobe. Further expansion will transfer matter from its envelope (unburnt hydrogen) onto the less massive star, and the first stage of mass transfer begins. Within just 50 000 years, almost $9 M_{\odot}$ are transferred onto the originally less massive star, now weighing in at around $17 M_{\odot}$! Only a $3.5 M_{\odot}$ helium core remains of the originally more massive component. However, conservation of mass and angular momentum have widened the binary orbit to just over 400 days.

The helium star (which we will see as a Wolf-Rayet star) does not survive for very long. Its energy source is now helium fusion (to form a carbon core), and within 2 million years, it will collapse and detonate as a supernova (the first one). The essentially instantaneous mass loss, combined with an asymmetry in the explosion, leaves a much longer period and highly eccentric orbit

(indeed it is at this stage that many binaries separate into single stars, accounting for the rarity of X-ray binaries).

The HMXB phase does not last for very long either, and once the mass donor evolves and expands, it forms a *common envelope* with the neutron star. The viscosity of the envelope leads to the neutron star spiralling in towards the massive star's helium core, and the orbital period shrinks to just a few hours. This dispels the outer envelope of the massive star, leaving a neutron star accreting by Roche-lobe overflow from the helium star. Evolution of the latter leads to the second supernova event, finally leaving a pair of neutron stars such as have been observed in the radio (e.g. PSR1913+16).

11.2.8 Extreme mass transfer in HMXBs: The microquasar SS433

Evolutionary calculations clearly show that HMXBs have very short lifetimes before the natural evolution of the mass donor leads to extremely high mass-transfer rates (as shown in Fig. 11.29) and the spiral-in of the compact object. Once thought to lead to the extinguishing of the X-ray source, and to be so rapid as to be unobservable, recent work suggests that the early stages of this entry into extremely high mass transfer could lead to exotic modes of behaviour.

One such object, SS433, is worthy of detailed study. It exhibits properties of both normal galactic X-ray binaries and the far more luminous radio galaxies and quasars. No, the SS does not stand for *strange star*, but it would have been appropriate if it did. In fact, SS stands for the initials of two

Evolution of an HMXB

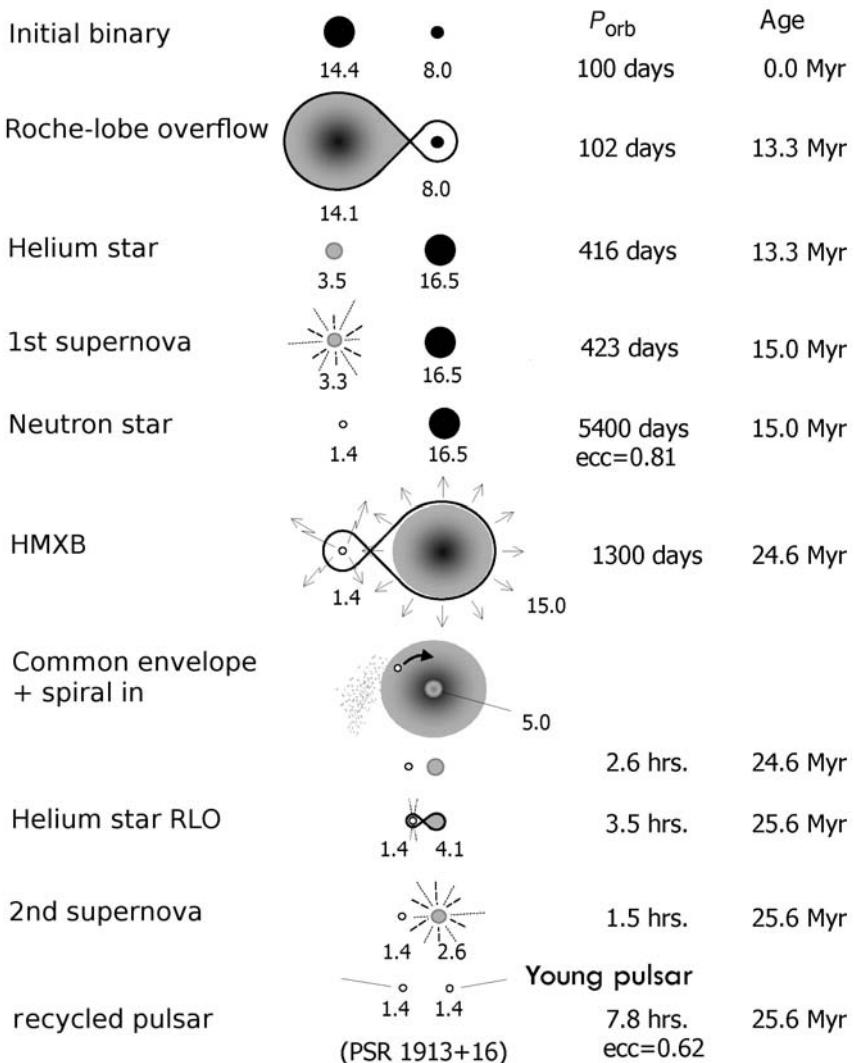


Fig 11.29 Evolutionary scenario for the formation of an HMXB/Be star from an initial binary of two stars of differing large mass at age 0. The more massive star evolves fastest, producing the first supernova event after 15 Myr. The resulting neutron star leads to the HMXB/Be phase once the surviving star has itself evolved to become a giant. This phase does not last long, and following a common-envelope ‘spiral-in’, a second supernova event can lead to the formation of a binary pulsar system (diagram based on Tauris & van den Heuvel, 2006).

astronomers, C. B. Stephenson and N. Sanduleak, who compiled a list of emission-line objects that was published in 1977. SS433 is the 433rd entry in that list.

The great majority of stars in our neighbourhood of the Galaxy are normal stars like the Sun,

whose optical spectra show a continuum with weak absorption lines that are characteristic of the elements in their atmospheres. However, unusual objects, such as X-ray binaries, cataclysmic variables, Wolf-Rayet stars and others, have totally different optical spectra. They tend to

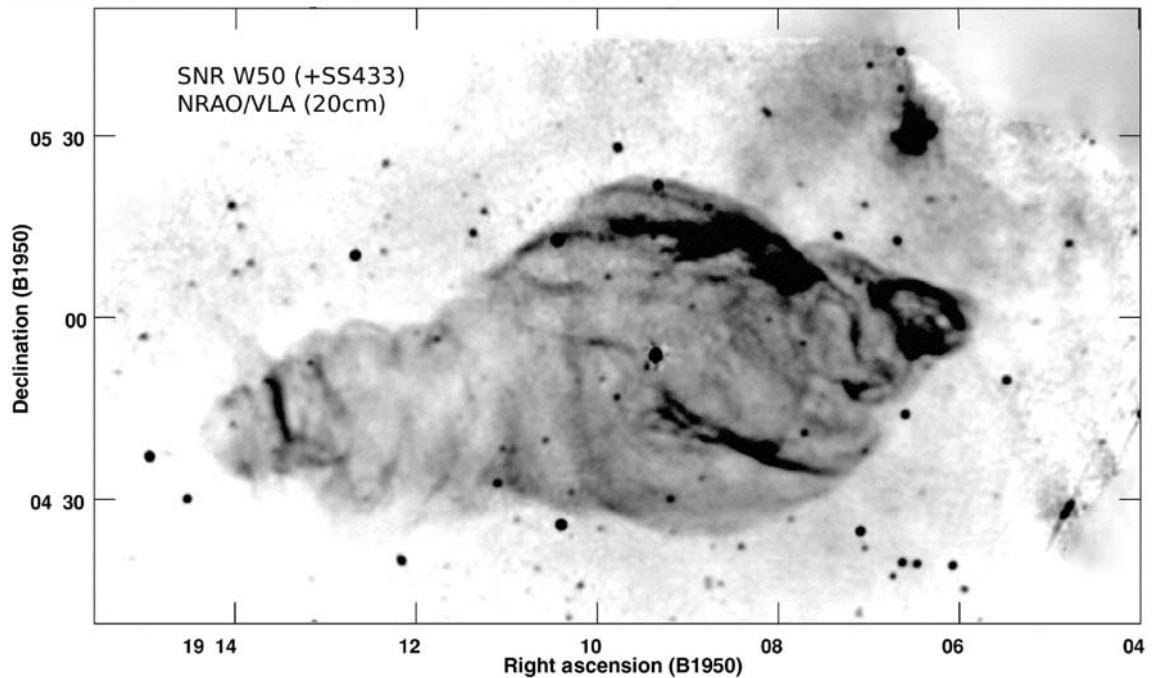


Fig 11.30 A radio map produced by the VLA at 20-cm showing SS433 itself as a central radio star surrounded by the shell of the supernova remnant W50. Note the extensions to the east and west of the essentially spherical central region of the SNR (image courtesy of NRAO/AUI).

be dominated by powerful emission lines, usually of hydrogen and helium, often indicating very high velocities due to mass in-fall and rotation.

Some stars are close to, or embedded in, cool material (matter that is neutral or, at least, not completely ionised). The result is usually the generation of emission lines. Energetic photons from the central source (particularly if it is a hot, early-type star) are absorbed in the surrounding material and excite the atoms to higher energy levels. Within a very short space of time (typically $\sim 10^{-8}$ s) such an atom returns to its original energy level, in the process emitting a photon of a characteristic wavelength. If this material is spatially removed from the exciting star (as seen from Earth), we see these photons as an emission-line nebula. HII regions or planetary nebulae are excellent examples, with the exciting star being a hot, massive star and a hot white dwarf, respectively. If the material is close to the star, there are emission lines seen in the stellar spectrum, which is usually an indication that something unusual is occurring.

11.2.8.1 How SS433 was (re-)discovered

The SS emission-line survey, including SS433, was published over a year before SS433's truly remarkable properties were recognised. Furthermore, a paper discussing several of the stars in the SS survey, again including SS433, was published 2 years earlier. Unfortunately, there was an error in the co-ordinates given for SS433, which would have prevented any observations of the object! Around the same time, the Cambridge 5-km radio telescope was completing the 4C survey for radio sources, which happened to include SS433. SS433 was in fact detected in this survey as a strong radio emitter, but remarkably, the published co-ordinates for it, too, were in error! The reason for the inaccuracy this time is illustrated in Fig. 11.30, a VLA radio map of the region of SS433. It shows an extended source of radio emission, rather like a spherical shell with ears attached on either side. More importantly, it shows the strong point source of radio waves at the centre of the shell. Unfortunately, the confusion in the output from the Cambridge interferometer caused by the surrounding diffuse shell of radio emission, resulted

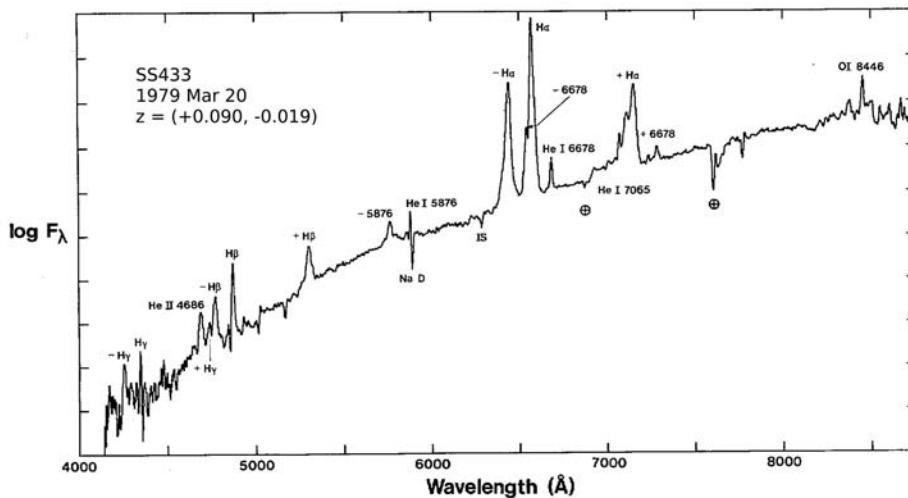


Fig 11.31 An optical spectrum of SS433 obtained from Lick Observatory with the 3-m Shane Telescope. The strongest features are those that made SS433 stand out in the emission-line survey, which is the spectral region around H α (diagram by Bruce Margon).

in the inaccuracy in the 4C survey position for the point source.

This radio shell is in fact W50, a supernova remnant first catalogued in the 1950s but then largely ignored. It was mapped again with greater accuracy in the mid-1970s, but the map covered only half of the remnant and thereby missed the fact that the point radio source associated with SS433 is in fact at the *centre* of the remnant. Had this been noted, the radio star would have been immediately thoroughly investigated as a candidate for the remnant of the core of the exploded star.

One of us (F.D.S.) was also involved in early observations of the SS433 region, this time at X-ray wavelengths in 1976 using Ariel V. He had thought that the X-ray source A1909+04 might be another extended source associated with the supernova remnant W50. However, his detailed analysis showed that the X-ray source was in fact variable and hence could not be associated with an extended object because of the very large light travel time across the remnant. F.D.S. and his colleagues also noted that the X-rays might originate from some compact remnant of the supernova explosion.

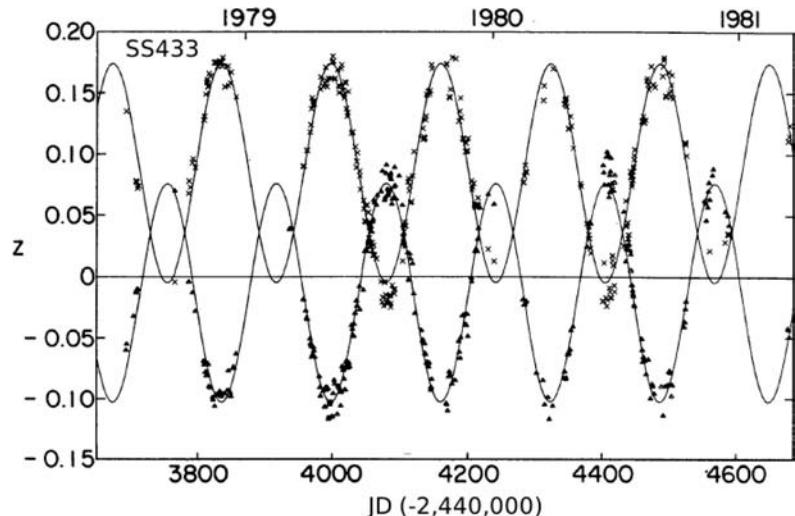
Finally, there was a deliberate search for optical stars coincident with point radio sources which might be stellar remnants of supernovae.

This was tedious because these mostly turned out to be quasars and not galactic objects. At this time, Martin Ryle and his colleagues realised that the strong radio source at the centre of W50 was at the same position as a 14th-magnitude star. They did not know that this star had already been catalogued by Stephenson and Sanduleak. Because of a variety of mistakes and coincidences, the peculiar properties of SS433 had been overlooked for several years after its initial discovery by the survey, but finally, the first optical spectra of SS433 were obtained by Clark and Murdin (1978) using the Anglo-Australian Telescope. Such is the way that scientific research actually proceeds!

11.2.8.2 The moving lines

These spectra of SS433 displayed strong emission lines, which was not surprising considering its presence in the SS catalogue. Clark and Murdin also found other emission lines at odd wavelengths that they could not identify but were the first to point out that the radio star, the X-ray source and SS433 were in fact the same object and that their location was at the centre of W50! This work came to the attention of University of California astronomer Bruce Margon, who began taking a series of optical spectra at Lick Observatory to study their temporal variations (Fig. 11.31).

Fig 11.32 A 4-year compilation of radial velocity measurements of SS433 against phase of the 164-day period, showing the excellent stability of the motion of the jets. The periodic motion, in opposite directions, of the blue and red moving lines is clear. The amplitude of the motion seen here is a staggering $40\,000 \text{ km s}^{-1}$ (diagram from Margon, 1984).



The principal Balmer line, $\text{H}\alpha$, appears to stay fixed throughout. However, there are two other emission lines, at atypical wavelengths, which, astonishingly, move from one night to the next. The velocity changes are staggering compared to any ‘normal’ stellar binary system. Remarkably, these extra lines, one to the blue of $\text{H}\alpha$, the other to the red, move smoothly back and forth through the spectrum. Margon discovered that the other emission lines have moving components, too. At certain times, the spectrum can become extremely complicated to interpret. Overall, though, each line’s set of moving components appears to follow the same pattern (Fig. 11.32), repeating their spectral motions smoothly about every 160 days.

Two obvious and quite remarkable properties are associated with this motion:

1. The amplitude of the movement is enormous. It implies a velocity of about $40\,000 \text{ km s}^{-1}$, far larger than anything ever seen in a galactic object. Indeed, it is quite easy to show from Kepler’s laws that such a motion, if due to two bodies orbiting each other with the 160-day period, implies that their total mass must be about 2 billion times that of the Sun! It is, of course, inconceivable that such a massive object within our Galaxy could have escaped detection hitherto.
2. The average, or mean, velocity of the motion is itself about $12\,000 \text{ km s}^{-1}$. Taken at face value,

it would imply, on the basis of the expansion of the Universe (i.e. employing Hubble’s constant), that SS433 is extremely distant and certainly extragalactic. However, the stationary emission lines are not at such a velocity, but are consistent with a very much lower velocity, as one would expect from an object located within the Milky Way.

11.2.8.3 Relativistic Jets

The physical origin of such fantastically high velocities emanating from a star that appeared to be in our Galaxy was a problem that provoked tremendous interest. The solution to this problem – now widely accepted as the basis for understanding SS433 – is the *kinematic model* developed independently by Fabian and Rees (1979) and Milgrom (1979), and fit to the optical characteristics of SS433 by Abell and Margon (1979). This model consists of a pair of oppositely directed jets of gas travelling at relativistic speed (Fig. 11.33). These jets precess about a common axis with the 164-day period, thus varying their orientation with respect to us and hence producing the observed Doppler velocity modulation.

It is not immediately obvious why the average speed of the jets at any instant and, in particular, their speed when travelling in a direction that is perpendicular to the observer is not zero. To understand this requires some knowledge of special relativity (see Box 11.6). If two clocks

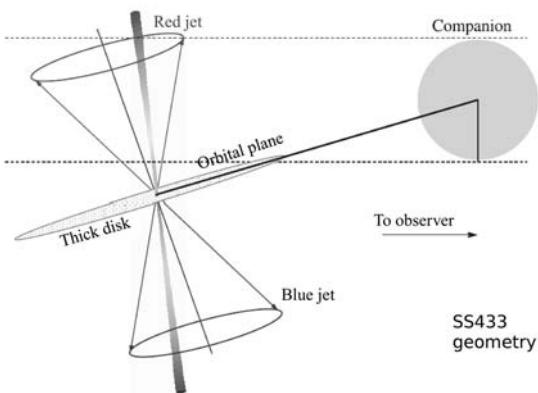


Fig 11.33 Schematic of the ‘kinematic model’ of SS433.

This shows the almost pencil-thin beams ejected in opposite directions at $0.26c$ along the rotation axis of an accretion disc which precesses with the 164-day period. Each beam or jet produces emission lines which we see as a *moving line*, due to the varying Doppler shift as the jet precesses. The angular extent of the precession is such that there is a time when the beams are travelling perpendicularly to our line of sight, and thereby the *moving lines* appear to cross over in the spectrum (Fig. 11.32).

are synchronised and then one is mounted on a spacecraft capable of very high speeds, some remarkable features emerge. An Earth-bound scientist will observe the spacecraft’s clock to be running slow compared to his own as it passes by him at high speed. This result is called *time dilation*, and it is independent of the direction in which the moving spacecraft is actually travelling. The clock appears to be running slow whether the spacecraft is coming towards you, away from you or travelling from side to side.

The notion of time dilation can be extended from the idea of moving clocks to the processes of atomic physics. Atoms are, after all, very precise little clocks. The light emitted by a moving atom will, because of time dilation, be seen by a stationary observer to have less energy (a longer, redder wavelength) than expected for that particular atomic transition. The independence of direction is most important as this is what distinguishes this effect from the *classical* Doppler shift of wave motions which depend in a very straightforward way on direction. A train whistle or police car siren will audibly change its frequency when travelling towards or away from the

Box 11.6 | The transverse Doppler shift

Consider an object travelling with a velocity v at an angle θ to our line of sight. The component of this velocity towards us is $v\cos\theta$. If v is small, then any radiation emitted by the object at wavelength λ_0 will be observed by us at a wavelength λ given by the classical Doppler shift:

$$\lambda = \lambda_0 \left(1 + \frac{v\cos\theta}{c} \right). \quad (11.19)$$

Note that this gives us the result we expect when $\theta = 90^\circ$, i.e. there is no Doppler shift when the object is travelling perpendicularly to our line of sight. The observed wavelength shift in a spectrum, defined by $z = (\lambda - \lambda_0)/\lambda_0$, is, of course, equal to $v\cos\theta/c$.

But when v is a large fraction of c , we must use a result from special relativity which gives the observed wavelength as

$$\lambda = \lambda_0 \frac{\left(1 + \frac{v\cos\theta}{c} \right)}{\left(1 - \frac{v^2}{c^2} \right)^{1/2}}. \quad (11.20)$$

This is the *relativistic Doppler shift*. A spectral line can be considered as an atomic clock which, at relativistic speeds, will appear to be running slow. The factor by which it does run slow is the $(1 - v^2/c^2)^{-1/2}$ term in equation (11.20) and is known as the *Lorentz gamma factor*. This effect is called *time dilation*.

Notice that even if $\theta = 90^\circ$, the observed wavelength is not the same as that emitted by the moving object. The observed wavelength is *always* greater than that emitted (i.e. it is red-shifted) regardless of the actual direction in which the object is travelling. In SS433 the relativistic time dilation causes the $12\,000\text{ km s}^{-1}$ average of the moving lines. From this, the true space velocity of the moving material is calculated to be $v = 0.26c$.

listener, whereas no change is heard if the vehicle is traversing across our field of view at some distance.

We can now see how to interpret the velocity curves of the two jets in SS433. The time-varying components are due to the normal Doppler shift

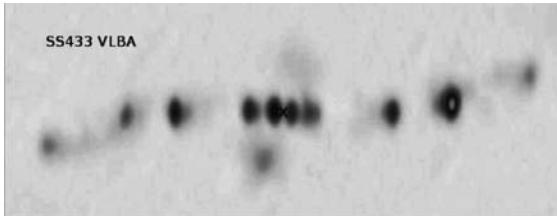


Fig 11.34 Global VLBA observations produce an ultra-high-resolution radio image of SS433 and its radio jets. Once material is ejected, it travels in a straight line, but the precession of the jets at the source produces a helix when projected on the plane of the sky. The field shown is only about 0.5 arcsec in width (image courtesy Amy Mioduszewski).

as the direction of motion of the jets changes (see Fig. 11.33). The *average* of the jets' velocities (which equals the cross-over velocity) is the time dilation effect, often called the *transverse Doppler shift*. Since it depends *only* on the absolute velocity of the jets and *not* their direction of motion, it can be used to determine the jet velocity directly (Box 11.6). It is $0.26c$, or $78\,000\text{ km s}^{-1}$. The relativistic nature of these jets is established beyond doubt.

11.2.8.4 Direct observation of the moving jets

The central radio 'star' of W50 (Fig. 11.30) was found by the VLA to consist of a highly variable point source surrounded by extended components which moved with time. These extended components were found to move outwards in opposite directions at a rate of 3 arcsec yr^{-1} . An ultra-high spatial resolution radio image of SS433 was obtained with the Very Long Baseline Array (VLBA) and resolved individual blobs of gas being ejected along the relativistic jets (Fig. 11.34). A movie showing the jet precessing through daily radio images can be seen online (<http://www.aoc.nrao.edu/mrupen/XRT/SS433/ss433plain.mpg>). By equating the spatial motion across the sky to the spectroscopically determined velocity of $0.26c$, and using the geometry of Fig. 11.33, the distance of SS433 can be determined. It is 5.5 kpc . What is more, the detail in this motion on the sky, when followed over several years, gives excellent support for the kinematic model.

The large amount of energy in these jets must surely have an observable effect on the large

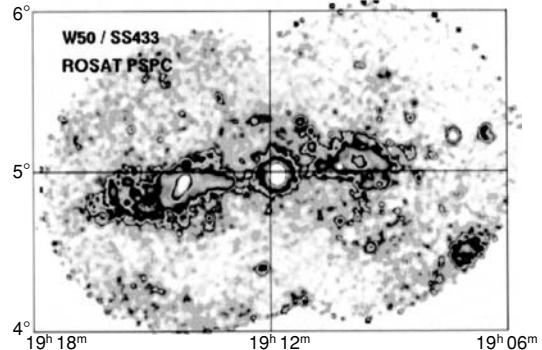


Fig 11.35 X-ray image of SS433 obtained with the ROSAT PSPC. The central point source is the accreting compact object of SS433 itself, whereas the X-ray lobes to the east and west delineate the precession axis of the jets (image courtesy NASA/HEASARC).

scale structure of the surrounding medium. With hindsight this can now be seen in the radio maps of W50 (Fig. 11.30), which show an essentially spherical shell with two ears or lobes that are aligned with the axis of the jets. Furthermore, the relativistically moving jet material must produce strong shock heating when colliding with the slower moving gas in the shell, thereby heating it to X-ray temperatures. This emission has been seen both close to the star and in the lobes of W50, as revealed in the ROSAT image of Fig. 11.35.

11.2.8.5 SS433 as an X-ray binary

So far we have concentrated on the dramatic properties of the moving lines in the spectrum of SS433 and the directly observable radio jets. The spectrum in Fig. 11.31, in addition to the lines from the jets, contains much more slowly moving features of an early-type B or A giant star. These features have a radial-velocity period of 13.1 days, confirming SS433 as an HMXB, albeit one with remarkable properties.

SS433 is clearly an interacting binary system involving very high mass loss from the massive donor star into an accretion disc surrounding a compact object. The inclination is high (it eclipses in both X-rays and optical), and the disc (or at least its inner regions) is precessing with respect to the orbital plane, thereby giving rise to the 164-day precession period of the jets. This yields the binary geometry shown in Fig. 11.33. The nature

of the compact object (i.e. whether a neutron star or black hole) is still a subject of debate. Observations of the mass donor are very difficult due to the bright, diffuse glow of the accretion disc. Nevertheless, the X-ray eclipse is long (~ 1.5 days), which places a lower limit on the size of the donor, as expected for a giant star.

The observed X-ray luminosity of the binary is surprisingly low, around 10^{36} ergs s $^{-1}$, but it is believed that the intrinsic luminosity is much higher. The high inclination of the binary orbit orients the rather thick disc so we never see the true or intrinsic X-ray luminosity. The X-ray eclipse is *partial*, so we probably only see scattered X-rays from the disc corona or wind, as in the case of the *accretion disc corona* systems amongst the LMXBs (see Section 11.3.3.7).

Estimates of the kinetic energy in the material of the moving jets, however, are almost 10 000 times higher at around 10^{40} ergs s $^{-1}$. This is based on the observed flux in the emission lines (optical and X-ray) from the jets. Such an energy release is not possible from a neutron star (see Box 11.7), and so the compact object is considered to be a black hole of at least $10 M_{\odot}$. An artist's impression of SS433 is shown as Colour Plate 30.

11.2.8.6 X-ray spectroscopy of the jets

An early discovery of the first generation of X-ray spectrometers (e.g. EXOSAT/GSPC, ASCA) was that the iron line (at 6.7 keV) in the X-ray spectrum of SS433 shared in the jet precession at the 164-day period. However, whilst its radial velocity variation was consistent with the kinematic model of the jets, only the blue-shifted component of most lines was detectable. This was explained by the geometry depicted in Fig. 11.36.

The emission line detected is that due to highly ionised iron which has lost all but one of its electrons and is thus at a high temperature ($50-100 \times 10^6$ K). It must originate very close to the central X-ray source. Chandra's high spectral resolution has allowed this to be investigated in much more detail (Fig. 11.37). This spectrum reveals a range of X-ray emission lines, corresponding to temperatures from 6 to over 100×10^6 K, and interestingly, a slightly higher jet velocity of 0.27c (implying that it originates from a physically distinct region, closer to the central source, than the

Box 11.7 | The Eddington limit

Matter falling onto a star or compact object encounters radiation from the star, which tends to retard it. The effect is usually very small, but the scattering of electromagnetic radiation off matter exerts a pressure on the material. This is proportional to the flux, F , of the incident radiation from the star. If the star has luminosity L and radius R , the radiation pressure acting on an electron produces a force of magnitude

$$-\frac{\sigma_0 F}{c} = \frac{\sigma_0 L}{4\pi R^2 c}, \quad (11.21)$$

where σ_0 is the Thomson cross section:

$$\sigma_0 = \frac{8\pi}{3} \left(\frac{e^2}{m_e c^2} \right)^2 = 6.65 \times 10^{-25} \text{ cm}^2 \quad (11.22)$$

It is possible, if L is large enough, for radiation pressure to overcome the gravitational potential of a compact object and thereby stop the accretion of matter onto it. The luminosity at which this happens is called the *Eddington luminosity*, L_{Edd} , and can be estimated from

$$\frac{\sigma_0 L_{\text{Edd}}}{4\pi R^2 c} = \frac{G M m_p}{R^2}, \quad (11.23)$$

which re-arranges to give

$$L_{\text{Edd}} = \frac{4\pi G M m_p c}{\sigma_0}. \quad (11.24)$$

The gravitational force of the compact object of mass M acts on a proton (mass m_p), and the protons and electrons are assumed to be electrostatically bound. Putting in the constants gives $L_{\text{Edd}} = 1.3 \times 10^{38} (M/M_{\odot})$ ergs s $^{-1}$, and so a neutron star of mass $1.5 M_{\odot}$ cannot produce a steady luminosity greater than about 2×10^{38} ergs s $^{-1}$.

In spite of the simplicity of this calculation (in particular, the assumption of spherical symmetry in the mass in-fall onto the compact object), this effect must be important in the generation of X-rays by neutron stars and black holes since there are few steady X-ray sources brighter than about 5×10^{38} ergs s $^{-1}$. In addition, there appears to be a maximum X-ray luminosity associated with the production of X-ray bursts on the surface of a neutron star which can also be attributed to the Eddington limit.

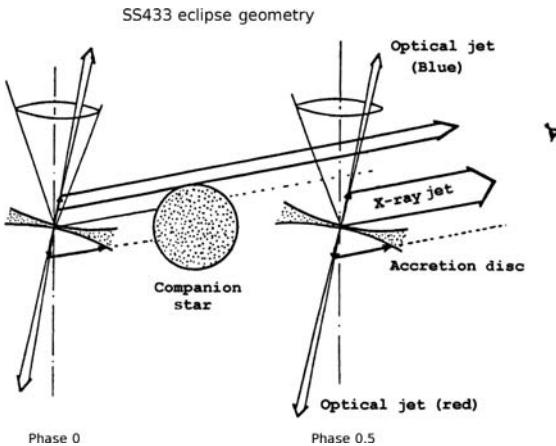


Fig 11.36 Geometry of SS433. The presence of only the blue-shifted X-ray component means that the red-shifted component is obscured from us by the accretion disc itself. The geometry needed to produce this effect is demonstrated here, with the accretion disc shown at opposite sides of the 13-day binary orbit. The upper X-ray jet will suffer a partial X-ray eclipse when on the far side of the companion at the same time that the optical emission from the disc is also eclipsed. The lower (receding) X-ray jet is never visible to us because it is always obscured by the accretion disc itself (based on an original by Masaru Matsuoka, ISAS).

optical jets). The narrow width of the lines also indicates a very narrow opening angle of only 1.2° (truly pencil thin).

A second Chandra HETGS spectrum obtained during eclipse revealed that the highest temperature lines of both jets were not eclipsed but that the cooler lines originating in the red jet were. The geometry of the emitting regions is then as indicated in Fig. 11.33 and supports the notion of decreasing X-ray temperature along each jet.

11.2.8.7 The importance of SS433

Since the discovery of SS433, the *microquasar* phenomenon has been seen in a number of other X-ray binaries (Chapter 12), most of which are black-hole X-ray transients. However, SS433 remains the prototypical and only example of the genre which displays precessing jets and is a *steady* and not transient source. Furthermore, it is the *only* jet source which emits spectral lines (hydrogen, helium and heavier elements), so that we know it is made of baryonic matter (i.e. heavy particles, such as protons and neutrons, as compared to the much

lighter free electrons and positrons). These moving lines give the precise velocity ($0.26c$) of the jet, and the X-ray lines give the temperature and density.

The analogy with AGN goes further with evidence for precession of their jets on extremely large scales. But these processes in AGN take enormously long timescales as well, whereas SS433's precession is 6 months, with detail in the radio jets observable on even shorter timescales. Since the physics is believed to scale quite simply with mass, SS433 provides a superb laboratory in which to study these phenomena.

The huge, and apparently steady, energy release raises its own questions. Theoretical models suggest that the mass transfer rate in SS433 is much higher than in any other known HMXB (as much as $10^{-4} M_{\odot} \text{ yr}^{-1}$) and that this will explain many of the peculiar features of SS433. In particular, there is a feature in the high-resolution radio image (Fig. 11.34) which consists of a faint, slightly diffuse ‘ruff’ at the position of the central source that is running almost perpendicular to the main jets. First noted by Blundell *et al.* (2001), this indicates the presence of a much slower, outflowing wind in the plane of the disc (see Colour Plate 30), almost as if there were a circumbinary envelope produced as a result of the huge mass transfer. Consequently, SS433 is now believed to be in the very short-lived phase of its evolution (Fig. 11.29), where the donor star is transferring material at such a high rate that it almost envelops the system. Such mass-transfer rates cannot be sustained for very long (a few tens of thousands of years only), and hence there will be few objects in such a phase at any given time. This would explain SS433's rarity.

As calculated in Box 11.7, the mass-transfer rate exceeds the Eddington value, which leads to large, radiation-driven mass loss from the central object. The likely flow of matter through the disc onto the compact object will be as sketched in Fig. 11.38, where the inner part of the accretion disc has so much material entering it that a substantial fraction is not accreted but is ejected along the rotation axis of the disc: the jets. This concept will return in Chapter 14 with the processes occurring in accretion onto supermassive black holes.

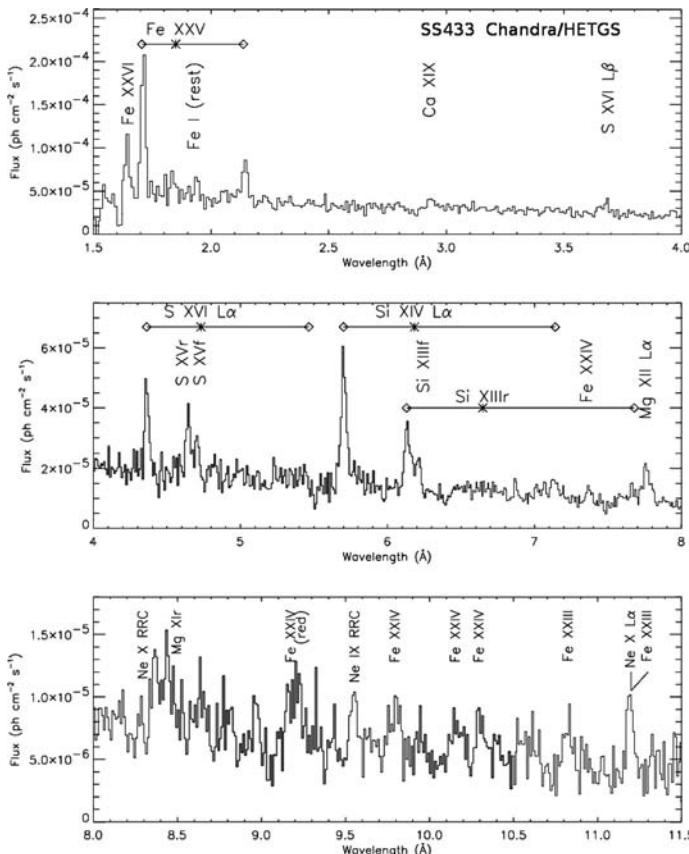


Fig 11.37 Chandra HETGS spectrum of SS433 obtained at precession phase 0.51 and orbital phase 0.67. Only emission lines from the blue jet are seen, except where noted. The strongest lines are marked with horizontal bars which connect the blue and red jet positions (diamonds) with their rest wavelengths (asterisks) (original diagram by Marshall *et al.*, 2002).

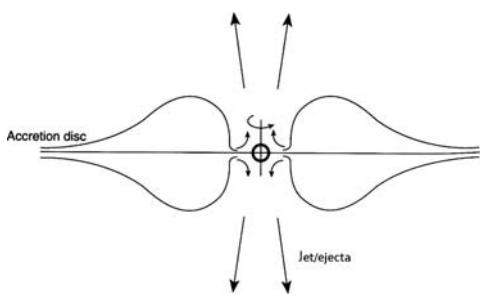


Fig 11.38 Schematic of how material flowing in the thickened inner part of the accretion disc can be constrained to be ejected along the disc's rotation axis. See also Chapter 14.

We are left then with this one glorious example of a star with precessing, relativistic jets. In spite of diligent observations over the last three decades of other stars in the SS catalogue and of many other X-ray binaries, no evidence for similar behaviour has yet been uncovered.

11.3 | Low-mass X-ray binaries

11.3.1 Fundamental properties

As the first extrasolar X-ray source to be discovered, and to this day the brightest X-ray object regularly visible, Sco X-1 is the prototype low-mass X-ray binary (LMXB). Since the optical spectrum (Fig. 11.1) showed no indication of the nature of the mass donor, it was not likely to be a wind accretor as in the HMXBs. Furthermore, the distribution of LMXBs around the Galaxy (Fig. 11.6) was concentrated in and around the centre of the Galaxy. They are associated with the older population II stars in the Galactic Bulge, and hence the mass donors for these X-ray sources must be comparable to or less massive than the Sun, as higher-mass ($> 1 M_{\odot}$) stars will have evolved and disappeared long ago.

The fundamental model for objects such as Sco X-1 was developed by Cambridge theoreticians Jim

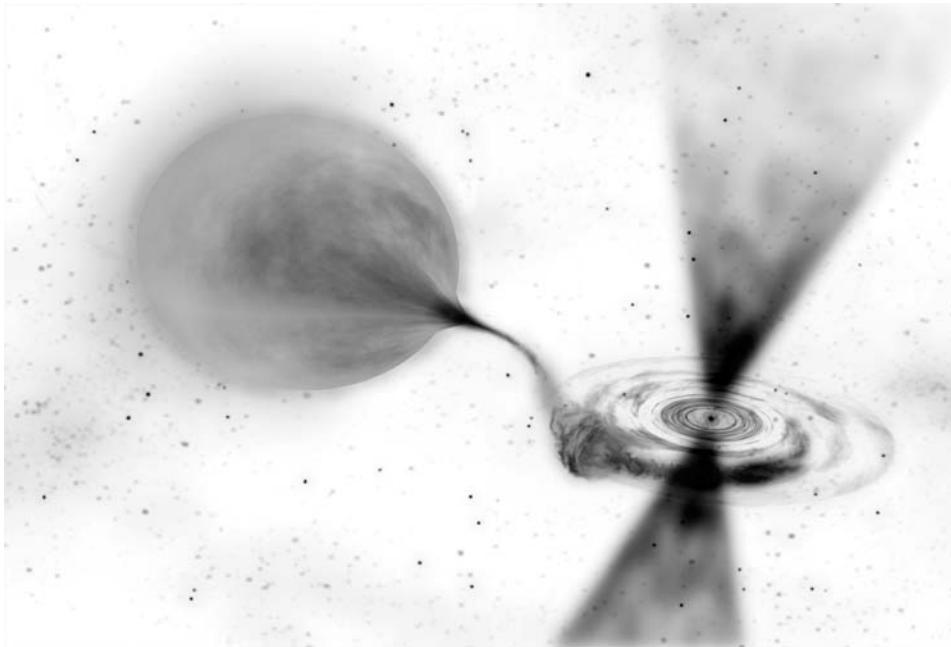


Fig 11.39 Basic model of an LMXB. The mass-losing star is of low mass and has evolved so that it fills its Roche lobe, thereby transferring mass that, because of its angular momentum, forms an accretion disc around the compact object, usually a neutron star. The accretion stream impacts the disc, forming a bulge which, if the inclination is appropriate, can occasionally obstruct our view of the X-ray source (producing dips). Frequently, there is evidence for jets of material from the inner disc, close to the compact object. Image credit: Dana Berry (SkyWorks Digital/NASA-GSFC).

Pringle and Martin Rees in the early 1970s. This basic idea (Fig. 11.39) will crop up again and again as different aspects are demonstrated by different sources. Essentially, a (slightly) evolved late-type star is filling its Roche lobe (see Box 11.2) and is transferring material through its inner Lagrangian point into the gravitational field of a compact object, which is either a white dwarf, neutron star or black hole. Because of the angular momentum of the material as it leaves the star, it cannot fall directly onto the compact object but must instead go into orbit around it, thereby forming an accretion disc. The formation and evolution of this disc forms one of the most important current branches of astronomy because it has applications that range from the formation of stars and planetary systems to galactic discs and to the active nuclei of quasars.

The major difference between this model and that for HMXBs is the mechanism of mass transfer from the normal star onto the compact object. A stellar wind carries very little angular momentum

and a (large) accretion disc is therefore unlikely to form. The wind from a low-mass (late-type) star, however, is feeble (as is the Sun's!) and would lead to completely undetectable levels of X-ray emission. Such a star must therefore direct its mass into the gravitational influence of the compact object to yield detectable X-rays. The intermediate step for this is through the accretion disc, formed by the angular momentum of the transferred material, and it is the disc which is responsible for much of the wide variety of behaviour seen in galactic X-ray binaries.

The key components of LMXBs as displayed in Fig. 11.39 are as follows:

1. *The compact object.* If the source is weak ($< 10^{33}$ ergs s $^{-1}$) and nearby (in the region of 100 pc), then the compact object is almost certainly a white dwarf. Above this luminosity, it has to be a neutron star or black hole. But if X-ray bursts, or pulsations, are seen, then it must be a neutron star. If it is luminous ($> 10^{37}$ ergs s $^{-1}$)

Table 11.7 Properties of LMXBs.

Source	P_{orb} (hours)	X-ray type	V	M_V	Notes
X1820-303	0.19	B, U, G	18.8	4.4	Likely triple
X1626-673	0.7	B, U	18.5		7.7-s pulsar, degenerate donor
X1916-053	0.83	B, D, U	21	5.3	X-ray and opt periods different, degenerate donor
J1808.4-3658	2.0	T, AMXP	16.6		
X1323-619	2.9	B, D			
X1636-536	3.8	B, A	17	1.3	
X0748-676	3.8	B, D, T	17	1.4	No decline
X1254-690	3.9	B, D	19		
X1728-169	4.2	A	17		~4-year superperiod
X1755-338	4.4	D	19		
X1735-444	4.6	B, A	17.5	2.2	
X2129+470	5.2	ADC	16.5		Now off; triple?
X1822-371	5.6	ADC	15.5		
X1658-298	7.2	B, D	18.3		
X1957+115	9.3		18.7		
Sco X-1	19.2	Z	12–14	0.0	Prototype LMXB
X1624-490	21	D			
2S0921-630	216	ADC	15.3		
Cyg X-2	235	B, Z	14.7	-2.0	Sub-giant companion

Note. See Liu *et al.* (2007) for a complete catalogue of LMXBs plus detailed references. B = burster; D = dipper; T = transient; U = ultra-compact binary; G = in globular cluster; AMXP = accreting millisecond X-ray pulsar; A = Atoll; ADC = accretion disc corona.

and has a soft X-ray spectrum, then it may be a black hole (see Chapter 12).

2. *The accretion disc.* Sources at high inclination enable the structure of the disc to be seen through its erratic absorption effects (dips) on the X-rays at certain binary phases. The inner part of the disc is very hot and is bright in the UV. If the inclination is high enough to prevent us from directly viewing the compact object at all, then these systems will only be detectable through their hot, X-ray-emitting coronae that extend above and below the disc.
3. *The secondary star, or mass donor.* Usually very faint and cool, they are impossible to see directly in most LMXBs. However, by searching in the IR, and by observing during eclipses, it is possible to show that the mass-losing star is of low mass (usually much less than the Sun) and likely evolved.

Table 11.7 lists some of the observed properties of key LMXBs that will be considered in this chapter.

11.3.2 Searching for binary behaviour

Establishing the validity of the model described earlier was not straightforward. Even showing that they were binaries was much harder than for the HMXBs. In fact, Sco X-1 and the other Bulge sources observed by Uhuru and other satellites in the 1970s showed no periodic modulation at all. Indeed, it took more than 10 years of careful study of Sco X-1 before the binary period was revealed, and even then it was from optical spectroscopy, not X-ray observations. By the mid-1970s, there were enough known sources so that it was realised by Milgrom (1978) that nature could not have placed us in just the right position that we would see no eclipsing sources. There had to be

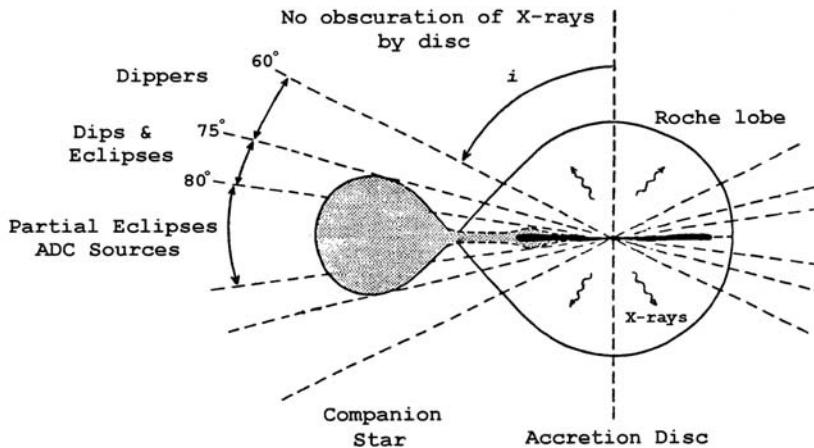


Fig 11.40 Schematic of the disc-shadowing concept. The accretion disc is thicker at the edge and protects the mass-losing star from direct irradiation by the X-ray source, as shown. An observer in the shadowed region will also not be able to see the X-ray source directly. This accounts for the absence of LMXBs that show observable total eclipses. When examined in more detail, the different viewing angles produce a variety of effects. At inclinations close to 90° (i.e. in the plane of the orbit and disc), the central source is not visible; only scattered X-rays from a wind or corona (ADC) can be seen (which are partially eclipsed by the companion star; see also Fig. 11.45). Between inclinations of about 60° and 80° the central source is visible, but with dips and other modulations superposed by the structure of the disc itself. Below 60° the central source is always visible, and very little in the way of orbital modulation can be detected.

a selection effect that prevented us from seeing high inclination systems. This is demonstrated in Fig. 11.40.

Viewing the LMXB in cross section shows how the disc itself is the source of the problem. An X-ray eclipse would be expected in cases where the mass donor comes between us and the compact object. However, it is at just these inclinations that the thickness of the disc itself prevents us from seeing any X-rays directly from the vicinity of the compact object. Such systems must exist, but we cannot find them by their X-ray emission. To be visible, the binary's inclination to us must be such that we can see over the disc edge down into the central region where the X-rays are produced. This is why early X-ray searches for binary behaviour were unsuccessful.

Much of the information about LMXBs contained in Table 11.7 has come from observations (usually photometry) of the optical counterparts. In LMXBs, the intrinsic faintness makes this possible (compared to HMXBs). The effects of X-ray heating (of the disc or companion star) are then more readily apparent. One of us (P.A.C.) has been active in this area, exploiting modern high-sensitivity

CCDs that enable accurate time-series photometry of faint stars to be performed. Orbital periods are now known for many LMXBs, and the relative sizes of a selection of these are shown in Fig. 11.41.

To put these into perspective, Fig. 11.41 also shows a schematic (to scale) of the 42-min binary 4U1626-67 superposed on an image of the Sun.

11.3.3 High-inclination binaries: X-ray dippers

There are other reasons why none of the low-mass systems showed binary evidence from early X-ray observations. Look again at Fig. 11.40 and imagine a system in which the companion star is just larger than the projected flare of the disc. If we were at just the right inclination, the top of the companion star would very briefly eclipse the compact object. In this case, the eclipse would occupy a very small fraction of the total binary cycle and hence be very difficult to catch. This proved to be true in the early days of X-ray astronomy because of the limited observing windows provided by satellites in low-Earth orbits. However, with the arrival of EXOSAT in 1983 (and the

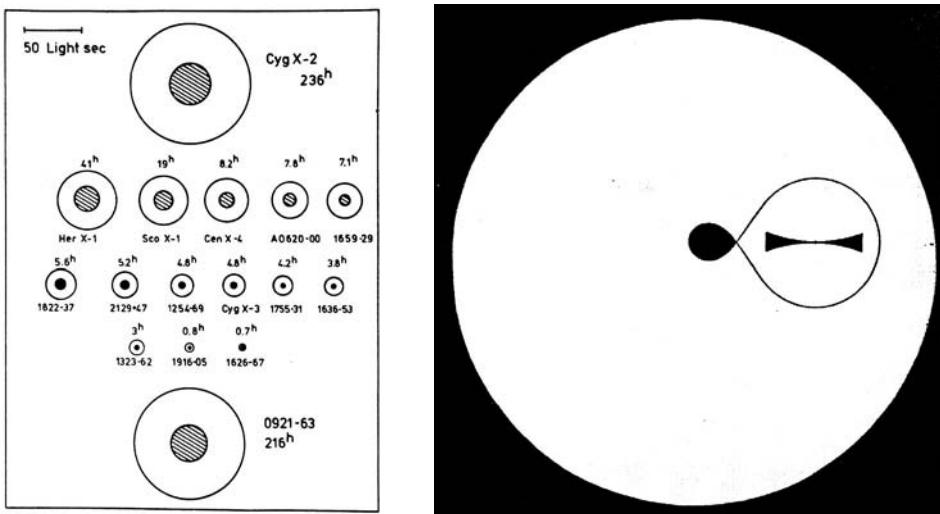


Fig 11.41 Relative sizes of the orbits and companion stars of a selection of LMXBs from (almost) the shortest to the longest periods. The bar at the top indicates the size corresponding to 50 light-seconds (the Earth is about 8 light-minutes from the Sun!). To put this into perspective, we also show the size of one of the LMXB orbits (MXB1626-67) relative to an image of the Sun (from diagrams originally by Saul Rappaport (MIT) and Jeff McClintock [Center for Astrophysics, CfA]).

more recent Chandra and XMM), it became possible to observe any X-ray source for long (approximately days), uninterrupted intervals. A new transient X-ray source, EXO0748-676, discovered by EXOSAT, demonstrated superbly the power of this extended monitoring capability. A continuous 20-hour observation is shown in Fig. 11.42, in which a wide variety of behaviour is contained in this one object. The most important feature is the eclipse, which lasts for just 8 minutes, out of a total binary period of almost 4 hours.

The upper part of this figure shows what the data would have looked like if they had been observed from a conventional, near-Earth-orbit satellite with 100-min period. The gaps destroy most of the fine temporal detail, and depending on the phase at which the observations start, it is possible to miss the eclipses altogether. The other main feature of this new source is the occurrence of X-ray bursts that permeate the light curve, which will be described later. These bursts indicate that the compact object in EXO0748-676 is a neutron star. Removing the bursts and then folding the data on the binary period of 3.8 hours gives the light curve shown in the lower part of Fig. 11.43.

An 8-min total eclipse is now readily visible. Its sharpness of ingress and egress indicates that the X-ray-emitting region must be very small, as expected for a neutron star compact object. However, this light curve shows other interesting features at phases between 0.6 and the eclipse. They look like erratic dips, and that is exactly what they are. They are a result of the line of sight to the X-ray source passing very close to the edge of the accretion disc (as demonstrated in Fig. 11.40).

To understand why they are confined to just part of the binary orbit, look back at Fig. 11.39. There is a region of the disc where the accretion stream from the donor impacts the disc and likely penetrates to smaller disc radii. Theoretical calculations suggest that the stream is thicker than the disc itself, thereby allowing it to easily penetrate to the inner regions of the disc. This causes the disc to have a non-uniform height around its periphery. When observing a source such as EXO 0748-676 at a high inclination, the variable height of the disc will obscure the neutron star from time to time. This idea is vindicated by examining the X-ray spectra of the dips. The amount of absorbing material (the column density) during the dip increases, exactly as expected. Because

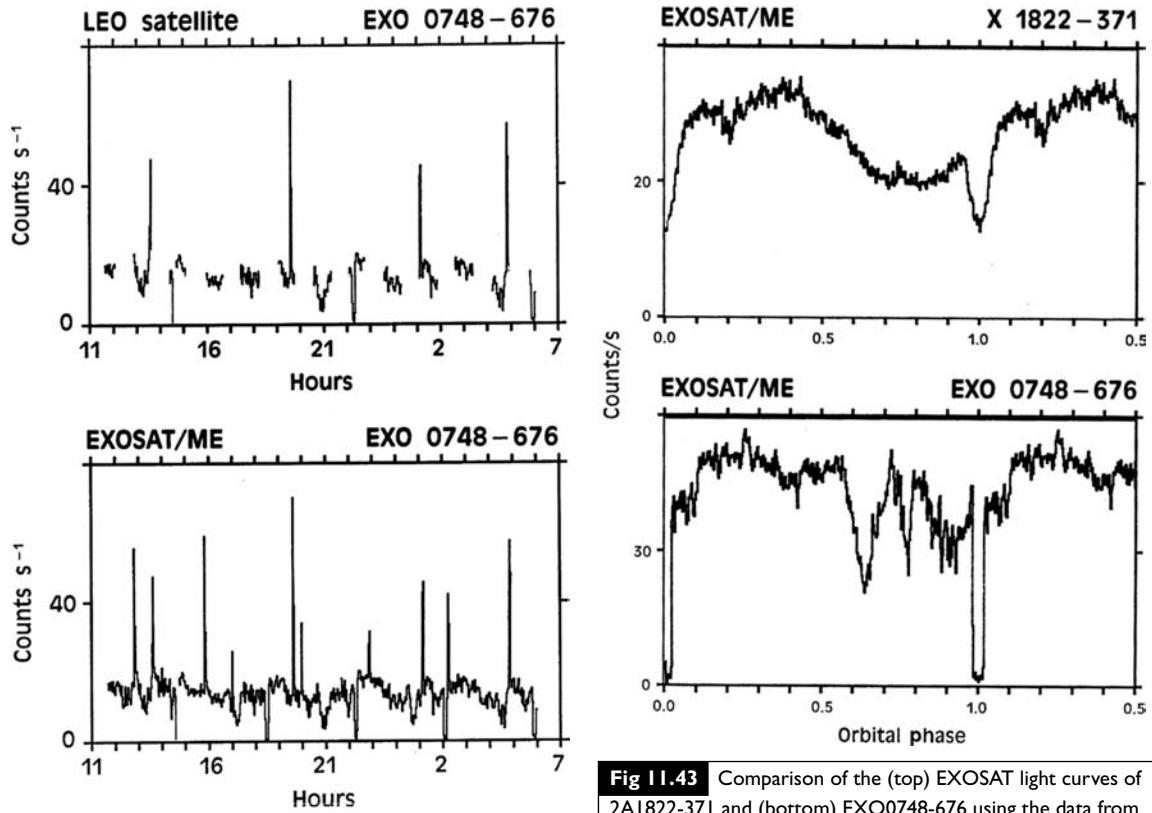


Fig 11.42 (bottom) A single 20-hour EXOSAT observation of the transient LMXB EXO0748-676. This source was discovered by EXOSAT during its manoeuvres around the sky (between other observations) and was found to exhibit dips, eclipses and bursts. (top) For comparison, this same light curve as it would have appeared to a conventional low-Earth-orbit satellite (courtesy of EXOSAT Observatory, ESA).

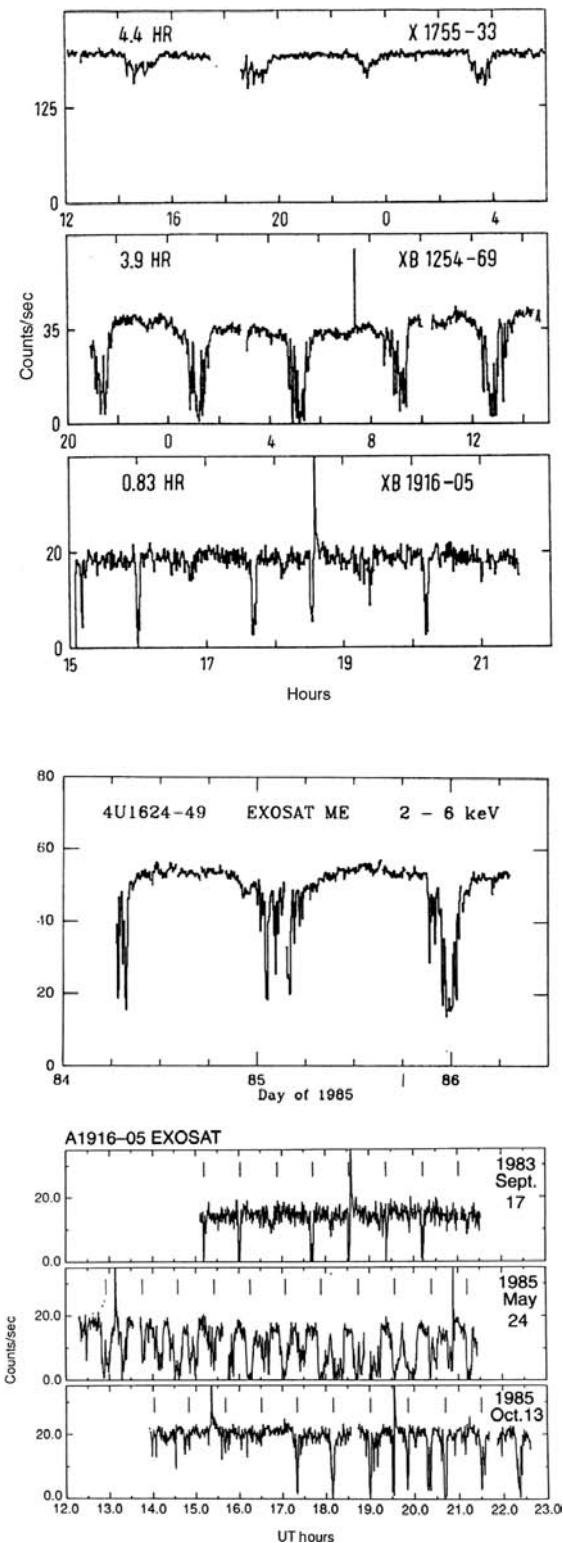
of the orbital motion of the two stars (and the fact that the mass donor is co-rotating), the mass-transfer stream will follow the path shown in Fig. 11.39. Since phase 0 is the time of X-ray eclipse by the companion star, the bulged part of the disc absorbs X-rays in the interval immediately preceding the eclipse.

X-ray dipping behaviour was discovered in Einstein observations of A1916-05 (Walter *et al.*, 1982). Because of its exceptionally short period (50 min), the regularity of the dips was recognised immediately (see Fig. 11.44). With EXOSAT, detailed study of this source was possible.

A1916-05 demonstrates very clearly that discs are not stable, rigid bodies, and the structure can

Fig 11.43 Comparison of the (top) EXOSAT light curves of 2A1822-371 and (bottom) EXO0748-676 using the data from Fig. 11.42 but with the X-ray bursts removed. Each shows 1.5 orbital cycles for clarity, with phase 0 defined as mid-eclipse; 2A1822-371 suffers a partial and broad X-ray eclipse combined with a smooth modulation, whereas EXO0748-676 exhibits a narrow but almost total eclipse combined with erratic variations. If these variations were smoothed out (as would be the case if the inclination were slightly higher), then the two light curves would be remarkably similar. The erratic variations are simply due to structure on the edge of the disc occasionally obstructing the view of the central X-ray source. In 2A1822-371, the central source is never visible at all (the inclination is higher), and so the eclipse and modulation are of an extended X-ray region, making the light curve appear much smoother (diagram courtesy of EXOSAT Observatory, ESA).

vary over long periods of time. Although the 50-min period is always visible (and is marked), the degree of obscuration presented by the disc edge can vary dramatically on a timescale of months. However, the dipping source most difficult for early satellite observations to have recognised was 4U1624-49. With its 21-hour binary period and quite spectacular light curve, its discoverers (Mike Watson and colleagues at Leicester) gave it the



name the ‘Big Dipper’! Note, however, the common feature of all these light curves, which is that there is no total eclipse present. In most cases, our line of sight to the X-ray-emitting object just misses the companion star but intercepts from time to time part of the edge of the accretion disc. The orbital inclination is very slightly lower than in the case of EXO 0748-676.

11.3.3.1 Accretion disc coronae

If the inclination is slightly higher (i.e. our line of sight is slightly closer to the orbital plane of the binary), then according to Fig. 11.40, we should see no X-rays at all. This is true if the X-ray-emitting region is confined to the immediate vicinity of the compact object. However, if there were an *extended* X-ray-emitting region, then it might be possible to see this if the luminous central source were permanently occulted. There is a small group of X-ray sources that display such behaviour, of which one of the best examples is 2A1822-371, whose light curve is displayed in the upper part of Fig. 11.43. There is an X-ray eclipse here, but it is *partial*. Note also the gradual ingress and egress as the companion star passes across our line of sight, the hallmark of an extended X-ray source. We cannot be observing the compact object itself, but the phasing of the partial eclipse tells us that the X-rays are in the same direction as the compact source. Hence this extended source must be

Fig 11.44 Montage of EXOSAT light curves of X-ray dippers. The orbital period is evident from the repetitive nature of the dips. The structure is more complex in some sources (e.g. XB1254-69 and 4U1624-49 compared to X1755-33), which may indicate slight differences in viewing angle (see Fig. 11.40). However, the disc structure in individual sources can evolve substantially with time, as shown in the observations of A1916-05, which are separated in time by many months. In the central panel, the disc has expanded to a height that creates substantial dipping structure for almost the entire orbital cycle. The very regular dips that are always present on the 50-min orbital period are marked with a series of vertical dashes. Note also that the X-ray bursts have been truncated here to make the dip structure stand out more clearly (diagrams based on EXOSAT Observatory originals).

Fig 11.45 Schematic of an accretion disc corona system. The neutron star is not visible to the observer, who only sees scattered X-rays from the extended corona above and below the disc. The bulge on the edge of the disc (where the accretion stream impact occurs) gives rise to a substantial modulation of the observed X-rays, and the mass donor partially eclipses the coronal region.

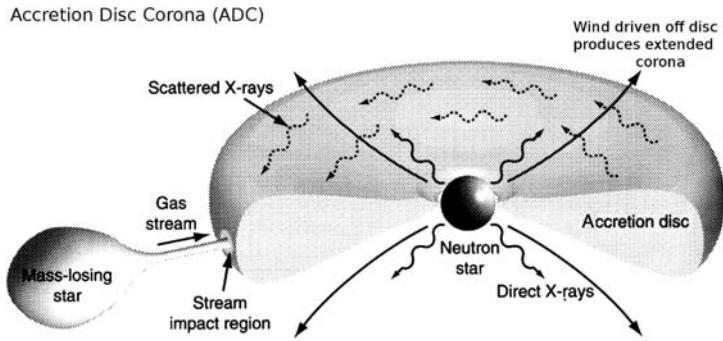
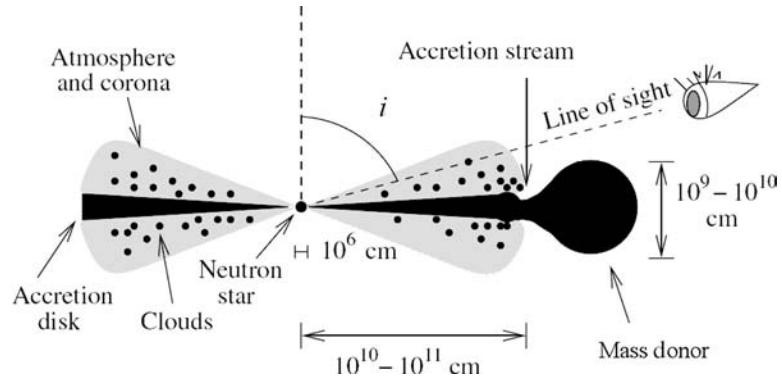


Fig 11.46 Observing a high-inclination LMXB means that our line of sight must pass just above the accretion disc and will reveal the regularly varying absorption due to material at the disc rim. The gas just above the disc has been driven off by X-ray heating from the central source, giving rise to a hot corona just above the disc proper (diagram from Jimenez-Garate et al., 2002).



a form of ‘corona’ which extends above and below the disc. We only see X-rays scattered into our line of sight from the accretion disc corona (or ADC) so the source is not strong. Indeed, 2A1822-371 is actually quite a weak X-ray source amongst the LMXBs.

The folded light curve of 2A1822-371 plotted immediately above that of EXO 0748-676 in Fig. 11.43 brings out their striking similarity. Admittedly, the dipping structure of the transient is very ragged, but if it were smoothed, it would follow the same general shape as that of 2A1822-371. Both are the result of a modulation of the X-ray source by obscuration due to the varying height of the edge of the disc. Because the X-ray source in EXO 0748-676 is pointlike, then the dipping behaviour will be highly erratic

since there is no reason why the disc edge itself should be smooth and well defined. However, if the X-ray-emitting region is extended, then the disc irregularities are averaged over the extent of the large X-ray source (see Fig. 11.45).

11.3.3.2 X-ray spectroscopy of the disc corona

LMXBs observed at high binary inclinations provide our best opportunity to study the properties and distribution of the gas associated with an X-ray-illuminated accretion disc. Those X-rays reach us through a sight line that passes just above the disc itself and are modulated by material along the disc rim (Fig. 11.46).

By observing X-ray dippers with Chandra and XMM, it has been possible to discern in detail the

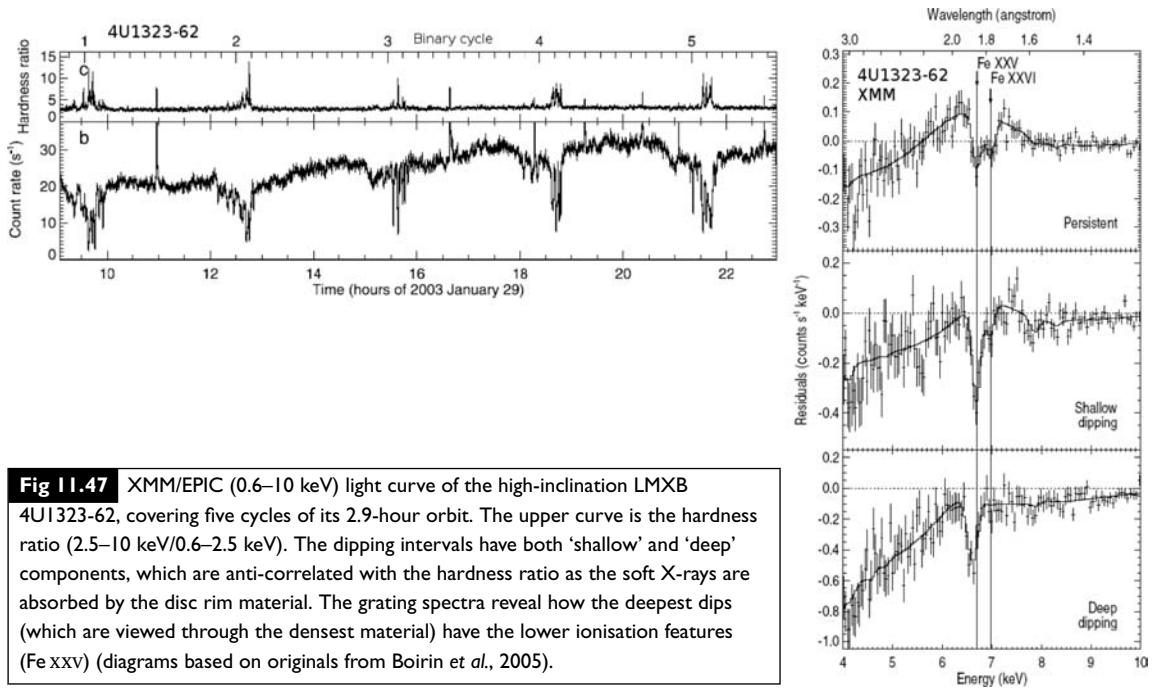


Fig 11.47 XMM/EPIC (0.6–10 keV) light curve of the high-inclination LMXB 4U1323-62, covering five cycles of its 2.9-hour orbit. The upper curve is the hardness ratio (2.5–10 keV/0.6–2.5 keV). The dipping intervals have both ‘shallow’ and ‘deep’ components, which are anti-correlated with the hardness ratio as the soft X-rays are absorbed by the disc rim material. The grating spectra reveal how the deepest dips (which are viewed through the densest material) have the lower ionisation features (Fe xxv) (diagrams based on originals from Boirin *et al.*, 2005).

characteristics of the gas that is producing the dips. An example is 4U1323-62 (Fig. 11.47), which was observed by XMM through both shallow and deep dipping, revealing highly ionised iron lines whose strength varies with the dipping. Comparing with spectra outside of dipping, it is clear that the most highly ionised (Fexxvi) gas is not visible during the deepest dipping, as expected when viewing through the densest (and least ionised) material. Similar behaviour has now been seen in all the dippers.

Almost all that has been learned so far about LMXBs has come from observations of periodic phenomena. The orbit of a binary adds the signature of a precise clock to the data, which immediately tells much about the object itself. In the mid-1970s and again in the mid-1980s, two types of X-ray variability were discovered which were to revolutionise our understanding of compact objects and the accretion process. These were X-ray bursts and quasiperiodic oscillations. Remarkable additional detail has come from the Rossi X-ray Timing Explorer (RXTE), which was designed specifically to study the most rapid X-ray variations.

11.3.4 X-ray bursters

Their very name conjures up an accurate picture of this X-ray event. First discovered in 1975 with ANS (Grindlay *et al.*, 1976; Belian *et al.*, 1976), the X-ray sources responsible were initially all located in globular clusters (see Section 11.4). Examples of X-ray bursts have already appeared in Fig. 11.42. These dramatic X-ray events are over in about 1 min and contain typically 10^{39} ergs of energy. (For comparison, it would take the Sun’s X-ray-emitting corona about 3000 years to radiate as much energy.)

The detailed properties of X-ray bursts are readily evident in recent observations made by RXTE/PCA. Its huge collecting area means that the signal-to-noise obtained is superb (see Fig. 11.48).

The start of the burst is extremely rapid, lasting about a second, and is simultaneous at all energies. However, the burst decay is very different in the various energy bands. At lower energies, corresponding to the coolest material, the burst displays a distinct tail, whereas it decays more sharply at higher energies. This means that the

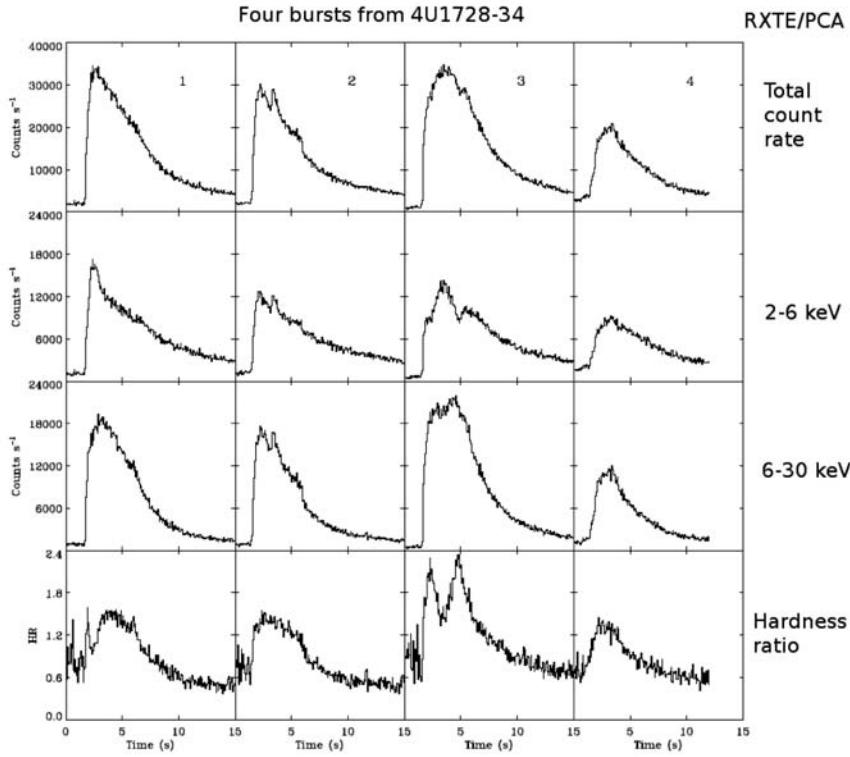


Fig 11.48 Four X-ray burst profiles from 4U1728-34 observed by the RXTE/PCA in different energy bands, as labelled. At lower energies the burst persists for much longer (it has a tail). This is because the radiating material cools significantly during the burst (based on original by Strohmayer & Bildsten, 2006).

temperature of the burst, which starts off high, falls as the burst progresses. It changes typically from about 30 million degrees at peak to about 15 million degrees some 10 s later.

While models involving a black hole were considered, it quickly became apparent that only a neutron star had the properties necessary to explain the observed details of X-ray bursts from a variety of X-ray sources.

11.3.4.1 Thermonuclear burning on neutron stars

From OSO-8 observations, Jean Swank and colleagues (at GSFC) noted that the X-ray spectrum of one particular burst was remarkably similar to that of a cooling blackbody. A blackbody is a perfect absorber and radiator (see Chapter 2) and can be described by just two parameters: temperature and radiating area. Apart from

being associated with globular clusters, the X-ray bursters were generally distributed in the Galactic Bulge, suggesting that bursting was an intrinsic property of Bulge sources and that they must be one and the same population. This infers a distance for most sources of 5–10 kpc. Armed with this assumption, Swank's analysis (see Box 11.8) showed that after an initial expansion, the radius of the (blackbody) sphere required to radiate the observed X-rays in the cooling tail was always in the range 10–15 km, exactly the expected size of a neutron star.

The burst process is a rapid one, and the accepted model involves unstable nuclear 'burning' on the surface of a neutron star, which leads to an explosive thermonuclear flash of radiation, with the peak energy occurring at X-ray wavelengths (Woosley & Taam, 1976). The first detailed numerical computations of such an event were

Box 11.8 Blackbody radiation from a neutron star

If the observed X-ray flux, F_X , during an X-ray burst from a source at distance d is due to blackbody radiation from a sphere of radius R and temperature T , then

$$4\pi d^2 F_X = 4\pi R^2 \sigma T^4, \quad (11.25)$$

where the right-hand side is merely Stefan's law. From this equation the radius can be calculated as

$$R = d/T^2 (F_X/\sigma)^{0.5}, \quad (11.26)$$

where everything, except the distance d , can be measured directly from the X-ray spectrum. However, regardless of whether d is known or not, it was clear from the observations that F_X/T^4 was a constant during the cooling tail of the burst and, for sensible estimates of d , was consistent with an R of about 10 km. The compact object must therefore be a neutron star, and the X-ray burst arises from physical processes occurring on its surface.

produced by Paul Joss of MIT and are illustrated in Figs. 11.49 and 11.50.

The accretion of hydrogen onto the neutron star from the accretion disc will produce a surface layer of hydrogen, which, in the extreme conditions prevailing, will burn (in the sense of nuclear fusion) steadily, producing an underlying layer of helium. Eventually, the density and temperature in the helium layer reaches a critical point, and it, too, starts burning, fusing into carbon. However, the helium burning process is inherently unstable, and all the helium will be rapidly consumed, producing a thermonuclear flash that we see as an X-ray burst. The steady accretion of fresh hydrogen from the disc then continues, and after a similar, but not exact, amount of time, another burst will occur.

This scenario ought to have an inherent timescale associated with it since, under approximately constant mass transfer rates, the time between bursts would also be approximately constant. Interestingly, only one LMXB, GS1826-24, appears to behave in this way and has garnered the nickname the 'clockwork' burster. It is an

Box 11.9 X-ray bursts as a 'standard candle'

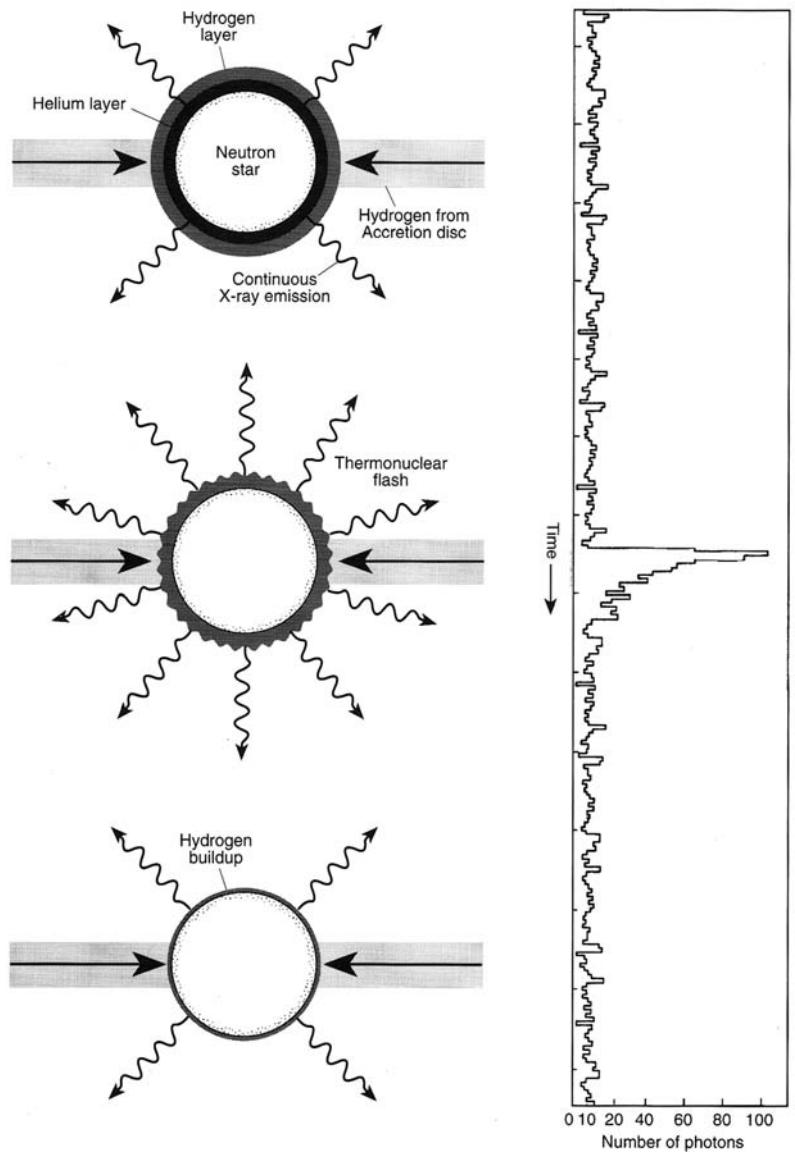
After a dozen or so X-ray bursters had been discovered, there appeared to be a maximum-burst luminosity that was never exceeded. This value was $\approx 1.8 \times 10^{38}$ ergs s $^{-1}$ and is exactly the Eddington limit expected for a $1.4 M_\odot$ neutron star, as described in Box 11.7. Since all neutron stars in binary systems seem to have about the same mass, the maximum burst flux seen from any burster is independent of distance and can be used as a *standard candle* to provide an estimate of the distance to these sources. This is a particularly valuable diagnostic for investigating the enigmatic Bulge sources and was used by Jan van Paradijs to estimate the distance to the centre of our Galaxy as about 6–7 kpc (assuming that the Bulge sources are uniformly distributed about the centre). This was significantly (30%) less than accepted values at that time (the late 1970s) but very close to the now accurately established distance to the Galactic Centre of 7.6 ± 0.3 kpc based on the orbits of stars very close to our Galaxy's central supermassive black hole.

important object for studying the physical processes occurring in thermonuclear flashes, and this is demonstrated in Fig. 11.51, in which the recurrence time between bursts from this object are plotted against the persistent ('steady') X-ray flux. The latter is a direct indicator of the rate at which matter is being accreted onto the neutron star, and if it increases, then burst frequency increases also because the fuel required is accumulating more rapidly.

This behaviour is as predicted by the thermonuclear flash model, although there are still details that we do not understand. Sometimes apparently normal bursts are followed a short time (only minutes) later by another burst, which is using fuel unburnt (for some reason) from the previous burst. This is believed to be related to non-uniform burning around the neutron star surface.

The length of time between bursts depends on the rate at which matter is being accreted, which, for typical values, is about 3 hours. Since the entire

Fig 11.49 Schematic of the thermonuclear flash model of an X-ray burst. (top) The neutron star is accreting hydrogen from its surrounding disc, forming a layer typically 1 m thick. This hydrogen fuses steadily into helium, forming a layer of comparable thickness. Eventually, the conditions in the helium layer go critical, and a thermonuclear runaway takes place (centre panel), producing the X-ray burst. The process then begins again (diagram by Walter Lewin, MIT).



process takes place within the surface layers of a neutron star, it is clear that the observation of bursts from an X-ray source rules out the possibility of the compact object being a black hole. Accreting material onto a black hole must pass through the event horizon and thereafter is unobservable. A surface is required on which the fusion can take place.

The thermonuclear flash model also accounts nicely for the observed ratio of the energy released in steady accretion to that released in X-ray bursts. An atom of hydrogen (actually, it would be a

proton because of the high temperatures) falling in the potential well of a neutron star would gain about 100 MeV of kinetic energy, which it would eventually release (mainly as X-rays) when it hits the neutron star surface. However, when that same atom (after a short wait) takes part in the nuclear fusion that we see as the burst, it releases only about 1 MeV of energy. We should see about 100 times more energy in steady emission (averaged over time) than in burst emission. This is exactly the ratio observed in many of the X-ray bursters. It also makes it clear that if the fusion

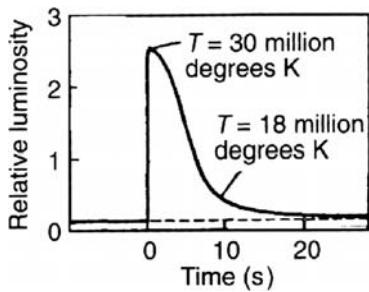


Fig 11.50 Theoretical model of an X-ray burst and its evolution with time. The model correctly reproduces the rapid rise, peak temperature and subsequent cooling (original diagram by Walter Lewin, MIT.)

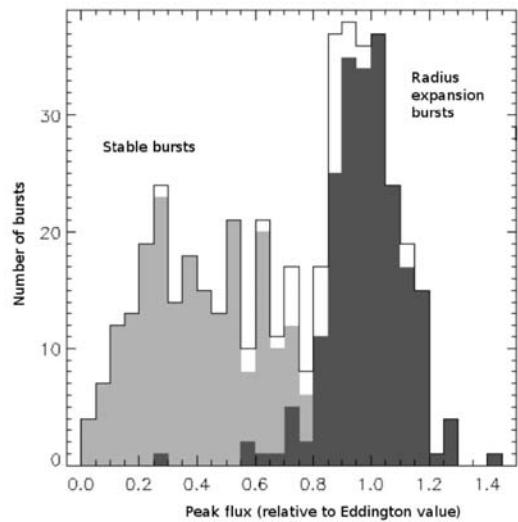


Fig 11.52 The distribution of peak fluxes of X-ray bursts observed with RXTE/PCA, expressed as a fraction of the Eddington limit. Those closest to the limit exhibit evidence of photospheric radius expansion (diagram based on original by Galloway et al., 2003).

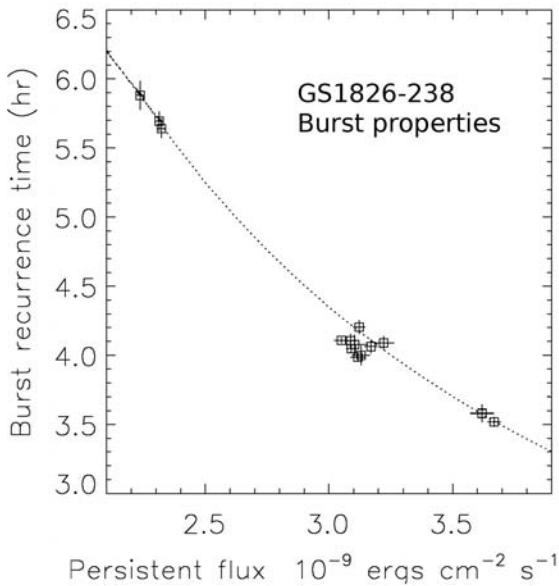


Fig 11.51 Relationship between the burst recurrence time and the persistent flux of GS1826-24 as observed by RXTE between 1997 and 2002. Clearly the recurrence time shortens as the persistent flux increases. This is as expected since a higher persistent flux implies a higher mass accretion rate onto the neutron star, and hence fuel for the burst is produced more rapidly (based on a diagram by Galloway et al., 2004.)

energy were released via steady burning (instead of these occasional bursts), we would never see it as it would be overwhelmed by the gravitational energy emission.

11.3.4.2 Photospheric radius expansion (PRE) bursts

As described in Box 11.8, the energy release in X-ray bursts is ultimately constrained by the Eddington limit. Figure 11.52 shows measured peak fluxes from several hundred bursts compared to the Eddington limit for a $1.4 M_{\odot}$ neutron star. Indeed, the peak flux is at, and does not exceed, this limit.

Consequently, we would expect the brightest bursts to exhibit properties showing that the huge radiation pressures reached actually succeed in lifting the top layers of the neutron star. When this happens, the effective temperature of those layers falls, simultaneously with an increase in the implied radius; i.e. there will be a photospheric radius expansion (PRE). The brightest bursts (which occur after the longest intervals, during which more helium fuel has built up) therefore show a double-peaked structure in the way that the temperature evolves during the burst. This is easiest to see in terms of the ratio of the hard to soft X-ray flux, as plotted in Fig. 11.48. Bursts 1 and 3 show this very clearly, indicating that the temperature drops as the radius expansion occurs, then recovers as it falls back. An analysis

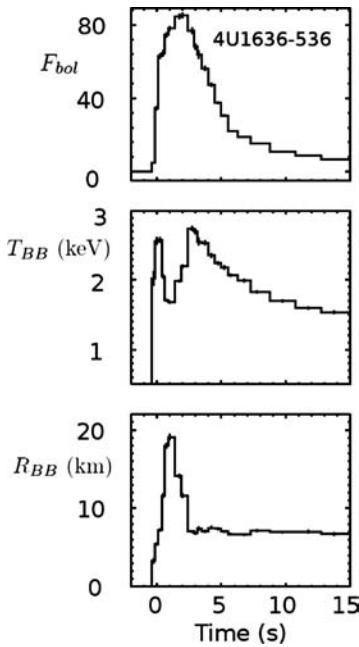


Fig 11.53 PRE burst exhibited by the prototypical X-ray burster, 4U1636-536, which shows how the blackbody temperature (middle) falls during the expansion of the neutron star photosphere (radius is given in the bottom panel). The temperature recovers as the material quickly falls back to the neutron star surface (diagram based on original by Galloway *et al.*, 2003).

of a PRE burst from 4U1636-536 is shown in Fig. 11.53.

Kuulkers *et al.* (2003) have shown that the luminosity at which this occurs for PRE bursts

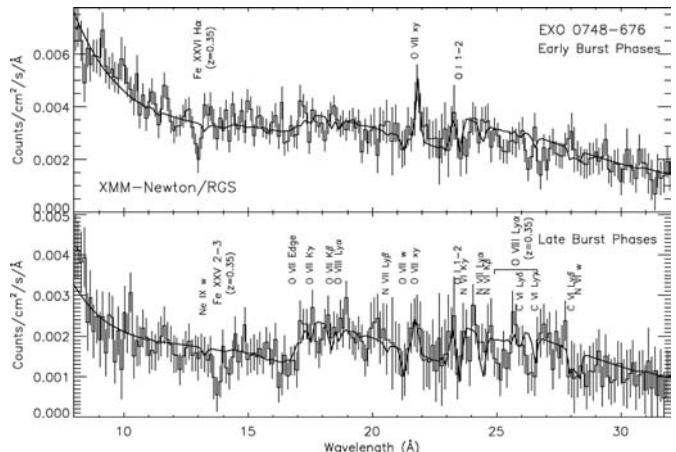
in a dozen globular cluster LMXBs (see next section) is 3.8×10^{38} ergs s $^{-1}$, with a range of only ± 15 per cent. These sources are almost certainly accreting hydrogen-poor material onto the neutron star, and the Eddington limit for helium is higher by a factor 1.75. The globular clusters present an ideal population in which to test these ideas because their distances (and hence luminosities) are much more accurately known than most other Galactic Bulge sources.

11.3.4.3 Spectroscopy of X-ray bursts

Obtaining direct evidence for such rapid motions during an X-ray burst, or of their detailed physical conditions, is difficult. The events are short-lived, and so spectral information is noisy. X-ray spectral features which have been gravitationally red shifted by the neutron star have been claimed over the last 15 years but never confirmed. Alternative explanations have been put forward (associated with the accretion flow and the irradiated disc), and the subject is controversial.

XMM provides the best opportunity because of its large collecting area combined with high spectral resolution. Extremely long observations of EXO 0748-676 have been undertaken, and Fig. 11.54 shows the XMM/RGS integrated spectrum of 28 bursts divided into their early and late stages. These spectra can be modelled well with features due to hot, ionised material in the vicinity of the neutron star. However, there are

Fig 11.54 XMM RGS-integrated spectra of 28 X-ray bursts from EXO 0748-676 during the (top) early and (bottom) late stages of the bursts. The best-fitting models of the X-ray spectra (solid curves) cannot account for the absorption features near (top) 13 Å and (bottom) 14 Å. These are interpreted as due to gravitationally red-shifted lines of (top) Fe XXVI (upper) and (bottom) Fe XXV (spectrum from Cottam *et al.*, 2002).



MXB1730-335, the rapid burster (SAS-3, 1976)

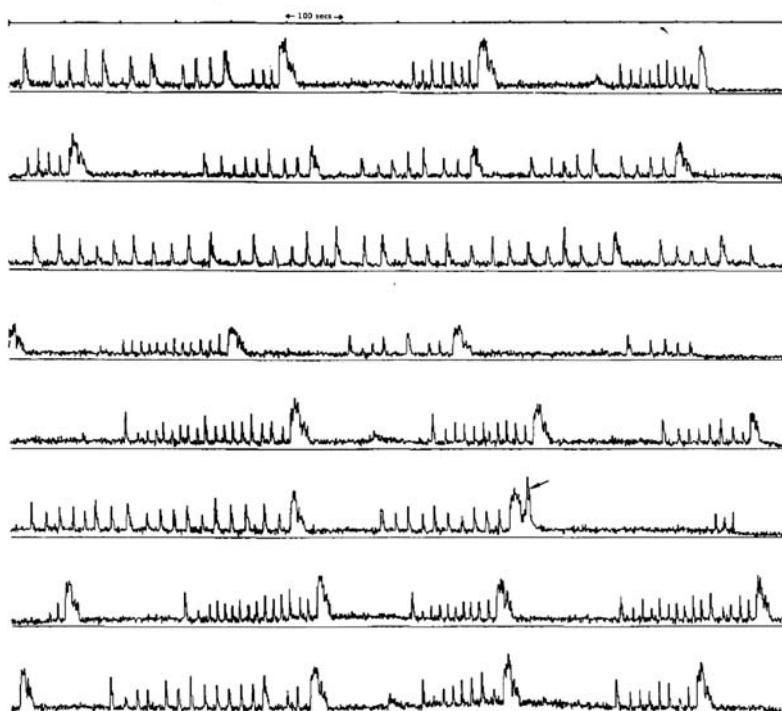


Fig 11.55 The Rapid Burster.

These eight orbits of SAS-3 observations (each panel covering 24 min) show its extraordinary bursting behaviour, which is classified as type II, together with a single, special burst, indicated by the small arrow, which is in fact an ordinary type I burst. Note how the smallest bursts are very close together, whereas a big one is always followed by a long gap (diagram by Walter Lewin, MIT).

particular unexplained absorption lines near 13 \AA and 14 \AA , which Cottam *et al.* (2002) suggest arise from Fe xxv and Fe xxvi and have been gravitationally red shifted by $z = 0.35$ (at the neutron star surface). Such a red shift would be expected for reasonable neutron star masses and equations of state. However, confirming and extending these observations will be a task for the next generation of X-ray telescopes.

11.3.4.4 The Rapid Burster

Early on in the study of X-ray bursters, an object was found that at first seemed to threaten the well-developing foundations of the thermonuclear flash model. This object was called the *Rapid Burster*, for reasons that are evident in Fig. 11.55.

These bursts look like astronomical machine-gun fire, recurring as rapidly as every 10 s. Sometimes there is a large burst that is followed by a longer gap until the next burst. The strongest bursts contain as much as 1000 times the energy of the weakest. This behaviour was at first a blow

to the thermonuclear fusion model because it cannot account for the behaviour of the Rapid Burster. However, on careful examination, it became clear that these rapid bursts had a very different character to those emitted from other X-ray bursters and therefore probably had a completely different origin.

Examine Fig. 11.55 carefully. The time between bursts is not random but depends directly on the strength of the previous burst. A big one is always followed by a long gap; a small one is followed rapidly by another burst. This is illustrated in Fig. 11.56, where the time to the next burst is plotted as a function of the energy in the burst.

Also, the X-ray spectra of these bursts showed no evidence for the cooling in the tail that is the hallmark of thermonuclear fusion; the temperature was essentially constant during the burst. As is usual in astronomy, when confronted with different types of behaviour, they were classified. The normal (or fusion) bursts are called type I, whereas the rapid repeating ones from the Rapid Burster are type II. A clue to the nature of the Rapid Burster is contained in Fig. 11.55. There is

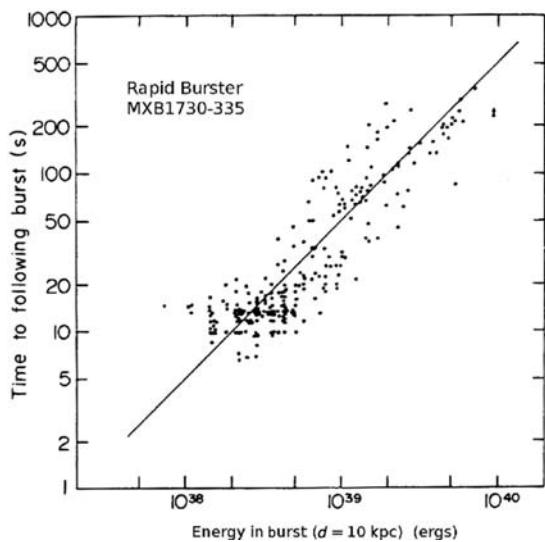


Fig 11.56 Burst properties of the Rapid Burster. As is qualitatively clear from Fig. 11.57, the time to the next burst depends simply on the energy contained in the previous burst (diagram by Walter Lewin, MIT).

one burst, quite a large one, that occurs at an unexpected time given the relation of Fig. 11.56. This burst is marked with an arrow and was called a special burst by its MIT discoverers Walter Lewin and Herman Marshall. On close examination, it was seen to be nothing more than a common or garden-variety type I burst. Even the Rapid Burster sometimes undergoes thermonuclear flash events.

This behaviour requires a neutron star which, in this case (see Fig. 11.57), has a strong enough magnetic field to produce a magnetosphere which acts as a gate that the accreting material has to pass through. Matter builds up around this barrier until it can no longer support the pressure. Suddenly, part of the magnetosphere gives way for a moment (the gate opens and then immediately closes), and matter can then fall directly onto the neutron star surface, producing an X-ray burst. If a large amount of matter manages to get through, then it will take some time for more to build up sufficient strength to force the gate (the magnetosphere) to open again. However, if only very little material gets through, then it is likely to open again within a short space of time. In physics this process is described as a *relaxation oscillator*. It accounts for the temporal behaviour of type II bursts very well indeed.

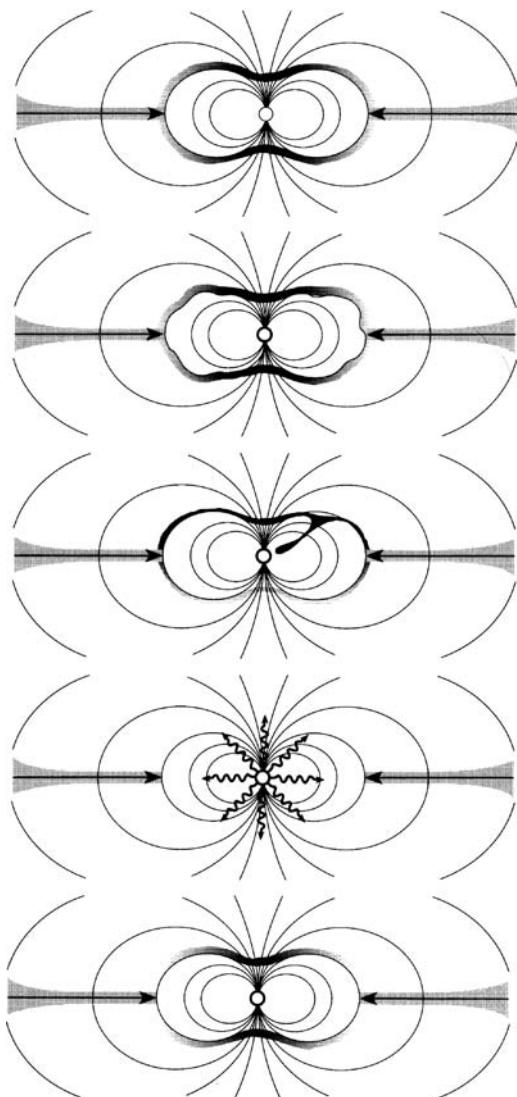


Fig 11.57 Magnetospheric gate model of the Rapid Burster. Material accreting from the disc is held back (top) by the neutron star's magnetosphere. When enough material has built up outside this gate, the magnetosphere can no longer hold it, and it ruptures (middle), thereby allowing it to fall onto the neutron star, producing a type II burst. With the material gone, the gate re-forms and the process starts again (diagram by Walter Lewin, MIT).

The accreted material is hydrogen that is passing through the gate onto the neutron star. If it were not for the magnetosphere, this would appear as steady accretion. This accreted hydrogen fuses steadily into helium, eventually producing a thermonuclear flash when the helium ignites,

which we observe as a ‘special’ burst. The special bursts are emitted every 3–4 hours, exactly as for other type I bursts. So after an initial scare, the thermonuclear flash model emerged intact from the Rapid Burster discovery, fully established as the basis for explaining X-ray bursts.

Remarkably, it was more than 20 years after the discovery of the Rapid Burster before GRO J1744-28 revealed itself as only the second known emitter of type II bursts. Known as the bursting pulsar, it consists of a 0.47-s X-ray pulsar in an 11.8-day orbit with its low-mass donor. The presence of strong pulsations and the absence of type I bursts indicate a stronger magnetic field than in the Rapid Burster.

11.3.4.5 Optical bursts too?

Within a year or two of the discovery of X-ray bursts, campaigns were organised to search for an optical counterpart. However, the technical difficulties were immense, and there were two problems that could not be avoided. Firstly, the sources were faint. The brightness of the steady optical counterpart of even the brightest X-ray burster had $V \sim 17$. Secondly, the exact times of X-ray bursts could not be predicted. Indeed, sources would often stop bursting altogether. This was compounded by, once again, the limitations of near-Earth-orbit satellites, whose observational coverage was usually only about 40 per cent. Add to this the vagaries of weather for Earth-bound observatories and the chances of success become slim, as was demonstrated by the complete failure to detect any optical bursts during the first season in which it was attempted (summer 1977).

The following year, the project moved to the Southern Hemisphere (Cerro Tololo), thereby gaining access to the brightest (optically) X-ray bursters, MXB1636-536 and MXB1735-44, and was successful in recording simultaneous optical and X-ray bursts for the first time. Nevertheless, the technology of those times was limited to the use of photomultiplier tubes for high-speed optical photometry, which greatly restricted the quality of the data on such faint objects. These bursts lasted only a few seconds, making them impossible to study with the first generation of optical CCDs. In the last decade, there has been a resurgence

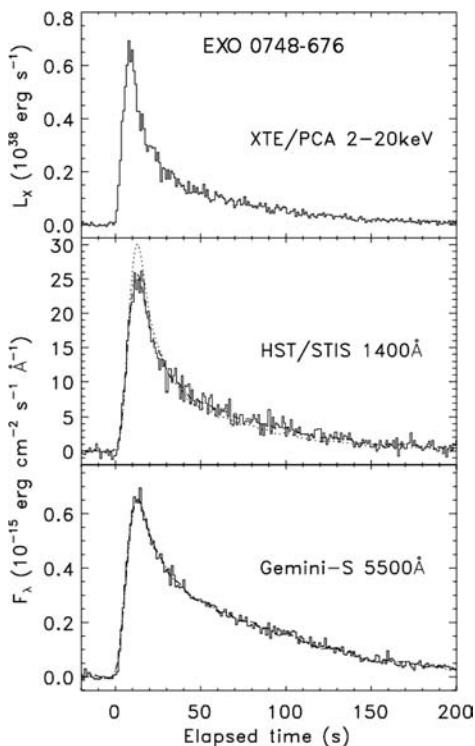


Fig 11.58 Simultaneous X-ray (RXTE/PCA), UV (HST/STIS) and optical (with the 8-m Gemini South telescope) bursts observed from EXO 0748-676 in February 2003. The X-ray burst profile is typical, but note the several second delay between the X-ray and UV/optical peaks. This is due to the time taken for the X-ray burst to irradiate and heat the disc, which then subsequently re-radiates this energy as the optical burst. The dashed curves represent model calculations of how the X-ray flux is subsequently absorbed and re-radiated by the accretion disc (diagram by Rob Hynes, Louisiana State University).

of interest in time-domain astrophysics, which has been able to exploit the development of fast-readout CCDs. It has even been possible to undertake high-speed UV studies of these events with HST, an example of which is shown in Fig. 11.58 which has simultaneous ground-based (Gemini South), UV (HST) and X-ray (RXTE) observations of a burst from EXO 0748-676.

The optical burst is delayed by a few seconds with respect to the X-ray burst, and for a given object, this delay varies little. Data show that the optical burst cannot therefore be caused by the X-ray burst heating up the surface of the companion star. If that were true, the delay would vary



Fig 11.59 Viewing geometry for EXO 0748-676 at the binary phase (0.38) corresponding to the X-ray burst observed in Fig. 11.57. The X-rays from the burst can irradiate the surface of the accretion disc plus the non-shadowed (shown with dark shading) areas of the inner face of the donor star (diagram by Rob Hynes, Louisiana State University).

as the stars orbit each other. Also, the level of smearing of the optical burst is much as would be expected if the X-ray burst were to heat up the surface of the accretion disc itself, which then re-radiates the energy at optical wavelengths (see Fig. 11.59). With data of this quality, Hynes *et al.* (2006) have constructed detailed models of the heated region of the disc which can reproduce the re-processed UV/optical emission. These are plotted in Fig. 11.58 and require the disc temperature to increase from a quiescent 18 500 K to a peak of 36 000 K. There is a 4-s delay between the X-ray and UV/optical peaks, but with considerable smearing (over almost 6 s) because of the physical extent of the disc.

Clearly the X-ray ‘flash’ and its reprocessed response at longer wavelengths has the potential to act as a diagnostic of the detailed geometry of the inner accretion disc surrounding neutron star LMXBs. If the response (i.e. UV/optical bursts) could be observed throughout the binary cycle, i.e. at different viewing angles, then we could infer the geometry directly. This analysis process is called *echo tomography* and is technically challenging given the inherent unpredictability of the X-ray burst mechanism. Nevertheless, the advent of very large ground-based telescopes, combined with a refurbished HST and the still-operating RXTE, augurs well for major advances in this area.

11.3.4.6 Superbursts

In the mid-1990s, the BeppoSAX mission contained an instrument, the Wide Field Camera

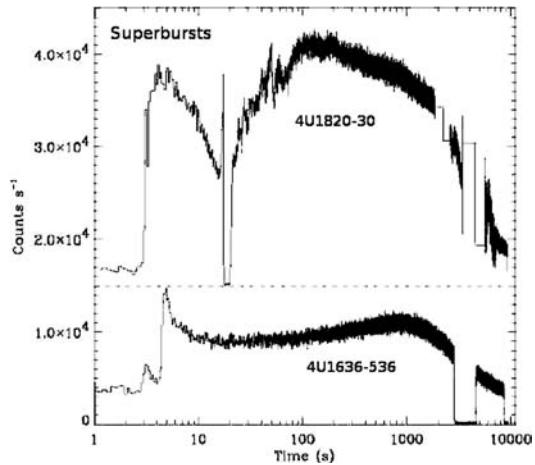


Fig 11.60 Two superbursts observed by the RXTE/PCA from the well-known X-ray bursters (top) 4U 1820-30 (with zero level set by the dashed line) and (bottom) 4U 1636-53. Note the logarithmic time axis, which allows a detailed view of the early superburst phase (diagram from Strohmayer & Bildsten, 2006).

(WFC), that revolutionised our view of the luminous X-ray sources in our Galaxy. While long (hours to days) observations of these sources had been accomplished with individual pointings (and still are), the WFC had an extraordinarily wide ($40^\circ \times 40^\circ$) field of view, which made it possible to monitor sources over long periods of time. In this way, it was discovered that a large fraction of the known X-ray bursters exhibited rare but extreme bursts that lasted several hours (Cornelisse *et al.*, 2000, 2002). Apart from their extreme duration (see Fig. 11.60), they otherwise had all the characteristic hallmarks of X-ray bursts, with a rapid rise followed by an extended, cooling tail. Consequently, explanations were sought which were related to the physics of ‘normal’ type I bursts, described in Section 11.3.4.1.

The duration (and hence greatly increased energetics) imply that the superbursts must originate in a deeper and hence larger volume of material. The end result of helium burning is carbon, and so the proposed culprit for superbursts is unstable carbon burning. However, this is a complex and still developing subject. Strohmayer and Bildsten (2006), for example, point out that the very detection of a superburst depends on

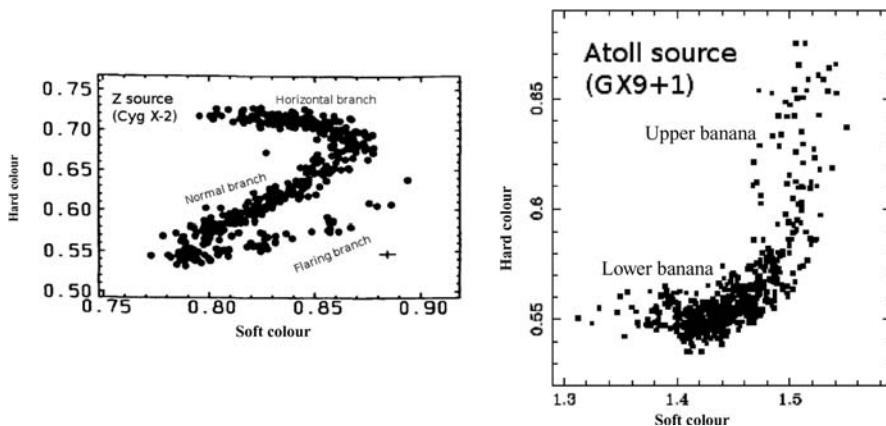


Fig 11.61 Broadband X-ray spectral variability in the form of ‘hard colour’ versus ‘soft colour’ for (left) the Z source Cygnus X-2 and (right) the atoll source GX9+1. Data points show ratios of fluxes in higher- and lower-energy X-ray bands, respectively, for short intervals of several observations. The names describe the shapes traced out by each source as they and their spectra vary with time. Within each of these, there are sub-divided labels corresponding to the ‘upper’ and ‘lower banana’ regions of the atoll and the ‘horizontal’, ‘normal’ and ‘flaring branches’ of the Z. The ‘flaring’ refers to the source’s short-term fluctuations (diagrams based on originals by Michiel van der Klis, Amsterdam).

the particular average flux level chosen for the burster itself.

11.3.5 Quasiperiodic oscillations

The neutron star in the majority of LMXBs is sufficiently old that its magnetic field is likely to have decayed significantly. These LMXBs are now classified into two main categories, ‘Z’ and ‘atoll’ sources, on the basis of their broadband X-ray spectral variability characteristics. The names are nothing more than a description of the shape that they trace out as they vary in a plot of their X-ray colour variations (Fig. 11.61, which also indicates some of the sub-divisions that occur within this categorisation).

We now know that this categorisation broadly links to the luminosity of the source (Fig. 11.62), with Z sources being Eddington limited and atoll sources typically 5–10 times lower in luminosity.

During its long lifetime as an LMXB, the neutron star should have been spun up by the continuous accretion of matter to a very rapid rotation, perhaps even close to breakup speed, which corresponds to a period of about 1 ms. However, the magnetic field would be weak, and any modulation at this period would be of very low amplitude

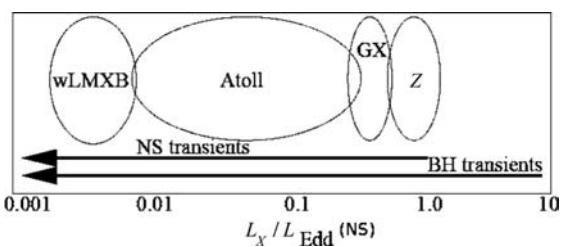
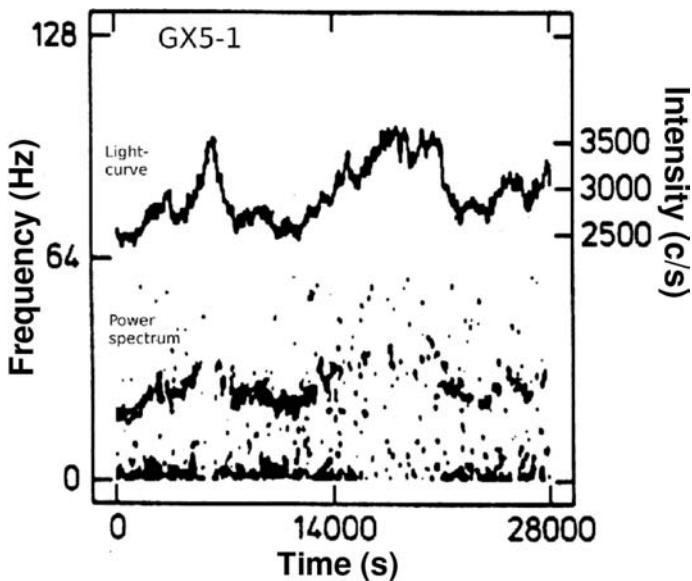


Fig 11.62 Range of X-ray luminosities (expressed as a fraction of the Eddington-limiting luminosity for a $1.4 M_{\odot}$ neutron star) for the different types of LMXB. The dominant LMXB types are the Z and atoll sources (the latter including those in the Galactic Bulge, labelled ‘GX’), and, together with the much fainter, low- L_X sources (‘wLMXBs’), they cover the same luminosity range displayed by the X-ray transients (diagram from van der Klis *et al.*, 1996).

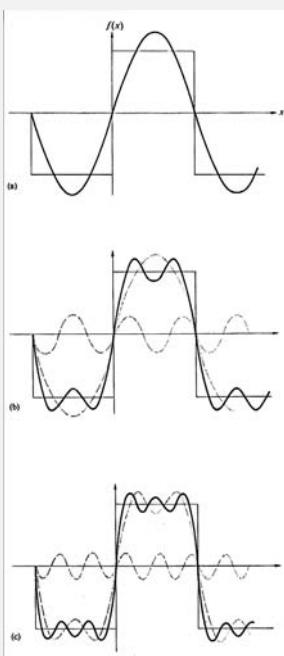
and difficult to detect. Large X-ray detectors would be needed to collect a significant number of X-rays within just a few milliseconds.

In the mid-1980s Michiel van der Klis and his colleagues at the EXOSAT Observatory observed one of the brightest of the Galactic Bulge sources, GX5-1 (so-named because of its position in galactic coordinates of 5° longitude and -1° latitude). The resulting light curve exhibited the remarkable property of *quasiperiodic oscillations*, or QPOs (Fig. 11.63).

Fig 11.63 The discovery of QPOs in GX5-1. The upper curve shows the X-ray light curve (right-hand scale) as seen by EXOSAT during 8 hours of observations of this very bright source near the galactic centre. Below this curve is the power spectrum of the same data (left-hand scale), which has been derived as a function of time during the observation. If there is an oscillation present in the data, it will show up as a peak in the power spectrum, e.g. at the beginning of the observation (time 0), there is an oscillation present at about 19 Hz. As the source varies in brightness, the frequency of the oscillations change, generally going to higher frequencies when it is brighter but disappearing altogether above a certain brightness (diagram courtesy of EXOSAT Observatory, ESA).



Box 11.10 | Fourier analysis and temporal variability



In the 1800s the French mathematician Joseph Fourier proved that any function (or series of data points) could be represented by just the sum of a set of sines and cosines. This is demonstrated in Fig. 11.64 using a square wave as the function to fit, and superposed is the sum of three appropriately chosen sine functions.

As more and more (smaller) terms are added together, a closer and closer approximation can be obtained to the square wave. In the limit of an infinite number of terms, the fit is exact.

Since a periodicity in data usually appears as something close to a sine wave, the dominant term in the Fourier series that represents the data will be that sine wave corresponding to the frequency of the periodicity. The noise, on the other hand, will be spread out over all frequencies, and thereby its effect will be greatly reduced. This is demonstrated using sample data in Fig. 11.65.

Fig 11.64 Synthesis of a square wave by the superposition of a Fourier series of sine waves. The pure sine wave at the top is a poor representation of the square wave, but the addition of each additional term (plotted in the centre) produces a closer approximation. If an infinite number of terms could be employed, then an exact square wave would be reproduced.

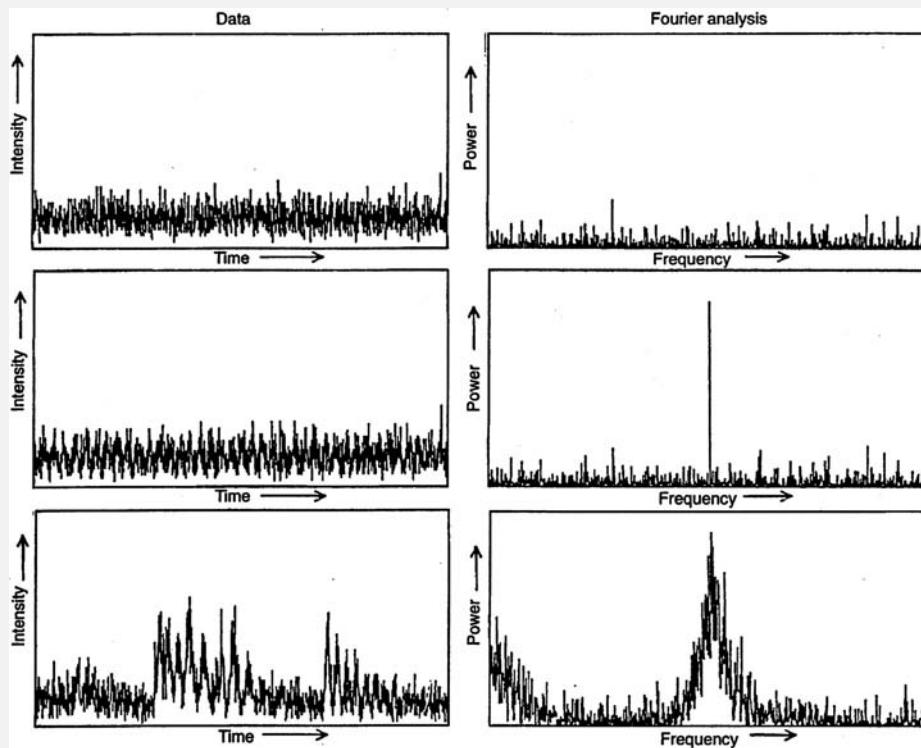


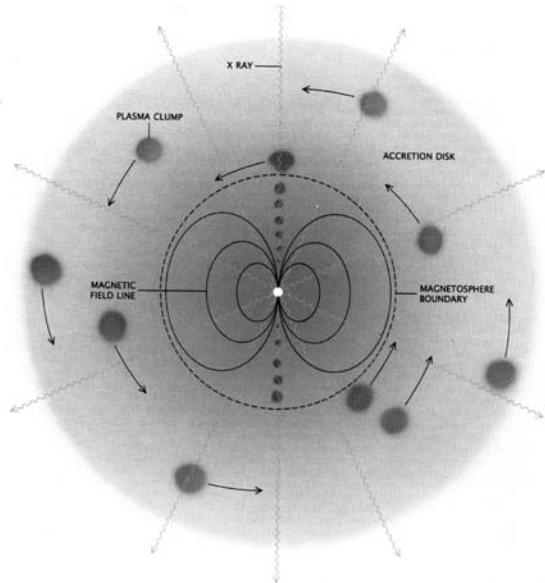
Fig 11.65 Demonstration of the ability of Fourier analysis to detect periodic signals in noisy data. There are three examples of such data here, shown in the left hand panels as counting rate against time. The top is pure noise, with no periodic signal present at all; the middle, although noisy, contains a regular pulsation; whereas the bottom data set contains pulsations that cover a range of frequencies (quasiperiodic). The Fourier analysis gives us information about how much power is present in the data at particular frequencies. Not surprisingly, the top power spectrum shows essentially a random scatter amongst all frequencies. However, the regular pulsation present in the middle panel stands out dramatically in this type of analysis (indeed, very much weaker periodic signals can be extracted than that in the data example here). The reason is that the noise in the data is uniformly distributed over all frequencies, whereas the regular pulsation is only present at one frequency, where its power is correspondingly greater. In the lower panel, the peak has been spread out because of the range of periods present in the data (diagram by Michiel van der Klis, Amsterdam).

The source was obviously varying rapidly, and this was investigated by Fourier analysis of the data (see Box 11.10, and in particular, Fig. 11.65, which illustrates the basic technique). This showed that there were regular variations present, but the period of the modulation changed with the intensity of the source. This effect can be seen in the figure, where the peak of the power spectrum is plotted below the source intensity curve. The period varies over the frequency range 20–40 Hz (i.e. periods of 50–25 ms), with

the frequency increasing as the source brightens. Because the period is not stable, but is always changing, it is called a *quasiperiodic* oscillation. However, at the brightest levels, the QPOs disappear completely.

This discovery was both exciting and puzzling. These were indeed rapid and semi-regular variations from an LMXB, but it was obvious that these oscillations could not directly represent the rotation of the neutron star because it is impossible to change the spin rate of such a massive object

Fig 11.66 The beat-frequency model of QPOs. This shows the inner part of the accretion disc surrounding the neutron star's magnetosphere. Both are rotating rapidly, but accretion of plasma clumps onto the neutron star can only take place when a clump finds itself over the pole of the neutron star. This occurs at a frequency which is the difference between the neutron star's spin frequency and the orbital frequency of the inner disc (based on a diagram by Michiel van der Klis, Amsterdam).



up and down so quickly. QPOs were also a factor 10 slower than had been expected for the actual spin rate.

11.3.5.1 The beat-frequency model

QPOs have been a subject of intense study over the 25 years since their discovery. Although slower than the anticipated spin period of a fully ‘spun-up’ neutron star, QPOs are still fast enough that they must originate from the inner regions of the accretion disc, where it interacts with the (weak) magnetosphere of the neutron star. Because of the weakness of the magnetic field, the disc will reach down almost all the way to the neutron star’s surface and therefore completely surround the magnetosphere (Fig. 11.66). Matter in these regions would be moving extremely rapidly just to remain in orbit around the compact object, as described in Box 11.11, and fast variations would be expected.

If the X-ray output increases, the accretion rate onto the neutron star must have increased. More matter flowing through the disc will lead to greater pressure on the magnetosphere, which will therefore shrink. Because it is then closer to the neutron star, the matter immediately

Box 11.11 Variability timescales close to the compact object

To understand how QPOs might arise, it is important to appreciate the typical timescales in the inner accretion disc region close to the compact object.

If matter were in free-fall from the inner disc, then its velocity v_{ff} would be $\sim(GM/R)^{1/2}$, which for a $1.4 M_\odot$ neutron star of radius $R \sim 10$ km would be $\sim 0.5c$. At such a speed, the matter would traverse 100 km in about 1 ms. Orbital periods for matter this close to the neutron star are similar. The orbital velocity at radius r is $(GM/r)^{1/2}$, which for $r = 100$ km corresponds to a period of approximately 15 ms. The inner disc matter must be moving very rapidly.

How close can the inner disc approach the neutron star surface and retain matter in a stable orbit? There are relativistic effects that come into play very close to the neutron star and produce an innermost stable orbit r_{ISCO} of $6R_g$, where R_g is the Schwarzschild (or gravitational) radius of $G M/c^2$. For a rotating $1.4 M_\odot$ neutron star, r_{ISCO} is ~ 12.5 km, for which P_{orb} is 0.6 ms.

above the shrunken magnetosphere will orbit faster, thereby producing oscillations at a higher frequency, as observed.

As so often happens, this beautiful explanation did not stand up to careful examination. The observed relationship between the QPO frequency and source intensity (as is clear for GX-1 in Fig. 11.63) could not be reproduced by the model. However, the basic idea of Fig. 11.59 was rescued by Lamb *et al.* (1985), who proposed that the X-ray modulation is not due directly to orbiting effects of the matter immediately outside the magnetosphere but instead is the result of the *difference* in frequency between this orbital period and the presumed more rapid spin period of the neutron star.

Returning to Fig. 11.66, imagine a large clump of matter orbiting just outside the magnetosphere, which is rotating rapidly beneath it at the neutron star spin period. The magnetosphere can be imagined as a barrier that will stop the clump accreting, unless it finds itself above the poles, which look like an open door. When this happens, part of the clump falls onto the neutron star, releasing a short flash of X-rays. It was shown that such a model could reproduce the spread of QPO frequencies and, more importantly, how they change with intensity. It is also necessary to assume that because of the weakness of the field (compared to HMXB pulsars), the hot spots on the neutron star surface (at the poles) are actually quite large and therefore smear out any periodic signal produced from the rapid spin. Otherwise, we would have expected to see the basic spin period itself in the data, too.

QPOs have now been found in virtually all bright LMXBs and Bulge sources, and as Michiel van der Klis pointed out at the time, QPOs could have been discovered much earlier. However, the astronomers were searching then for a strictly *periodic* signal, not a broad peak in the power spectrum. Unless a sufficiently wide range of frequencies is displayed in the power spectrum, it is quite easy for the broad peak to go unnoticed. This is demonstrated very nicely in Fig. 11.67, which also shows how much the power spectra vary with the spectral state (see van der Klis, 2005, for a much more extensive review).

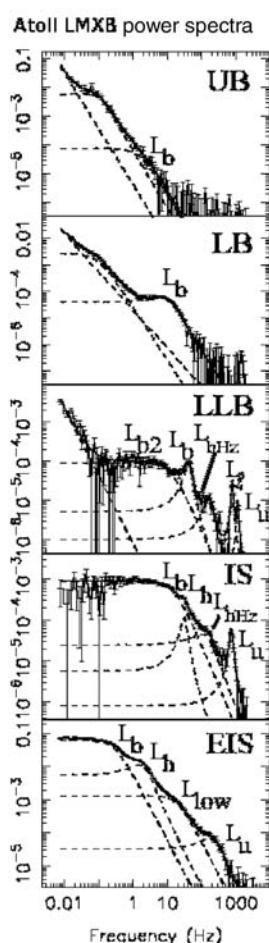


Fig 11.67 Power spectra of atoll sources in different spectral states (different ‘regions’ of the atoll, with LB and UB referring to the lower and upper banana of Fig. 11.61), showing the range of QPO phenomena observed. Note the breadth of some of these peaks, making them very difficult to discern (based on an original diagram by Michiel van der Klis, Amsterdam).

11.3.5.2 Discovery of kilohertz QPOs

The first decade of QPO studies was brought to a dramatic close with the launch of RXTE and the almost immediate discovery of the long-sought, ultra-rapid variations from luminous LMXBs. Figure 11.68 shows the remarkable twin-peak QPOs in the kilohertz regime, which required the ultra-high time resolution of the RXTE/PCA instrument and were first seen in the atoll LMXB 4U1728-34 (Strohmayer *et al.*, 1996) and then the most luminous LMXB of them all, Sco X-1 (van der Klis *et al.*, 1996).

These twin kilohertz peaks display very interesting properties. They are not stable but move in tandem, up and down in frequency, changing with the spectral state and luminosity of the source. The peaks are usually separated by ~ 300 Hz, but this can decrease slightly as their

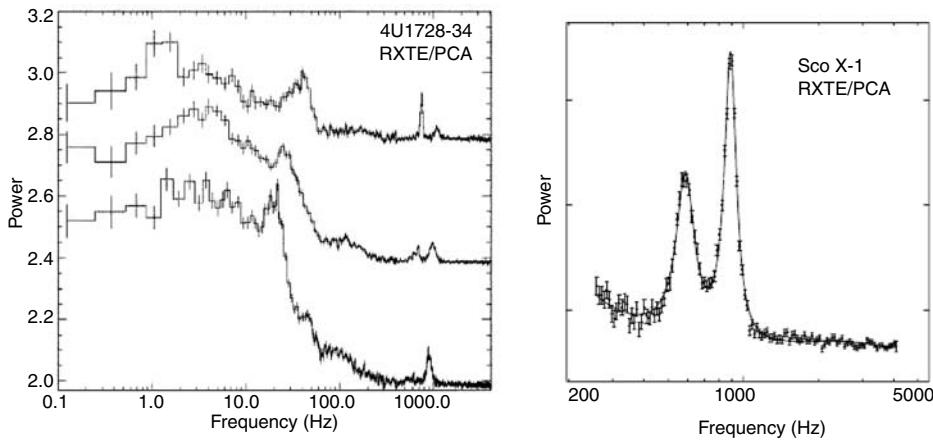


Fig 11.68 The large detecting area and high temporal resolution of the RXTE/PCA led directly to the discovery of pairs of kilohertz QPOs from (left) 4U1728-34 (Strohmayer et al., 1996) and (right) Sco X-1 (van der Klis et al., 1996). Several observations of 4U1728-34 are shown, demonstrating the volatility of QPOs as well as the presence of the slower QPOs (around 10–50 Hz) that were already known in atoll sources (diagrams by Michiel van der Klis, Amsterdam).

overall frequency increases. As described in Box 11.11, these frequencies are close to the periods expected for matter orbiting near the compact object. They indicate the presence of a ‘preferred’ radius in the inner disc, which would imply that limits can be set on the mass and radius of the neutron star since such a radius must be $>r_{\text{ISCO}}$, the innermost stable circular orbit. The maximum frequencies seen so far in kilohertz QPOs are ~ 850 – 1330 Hz, and if the latter does correspond to r_{ISCO} , then it requires a $\sim 2 M_{\odot}$ neutron star (Fig. 11.69).

Where known (see the next section), the peak separation is close to the spin frequency of the neutron star. This leads naturally to a beat-frequency explanation, in which the higher-frequency peak (located at a specific, possibly resonant, radius in the disc) beats with the neutron star’s spin to produce the lower-frequency peak. However, there are clearly complications in this interpretation as there appears to be a split at ~ 400 Hz, where the peak separation changes from representing the neutron star’s spin frequency to being just half that value (Fig. 11.70). QPOs have also been detected in black hole LMXBs (Box 11.12 and Chapter 12), but with only a single peak due to the absence of a spin signal.

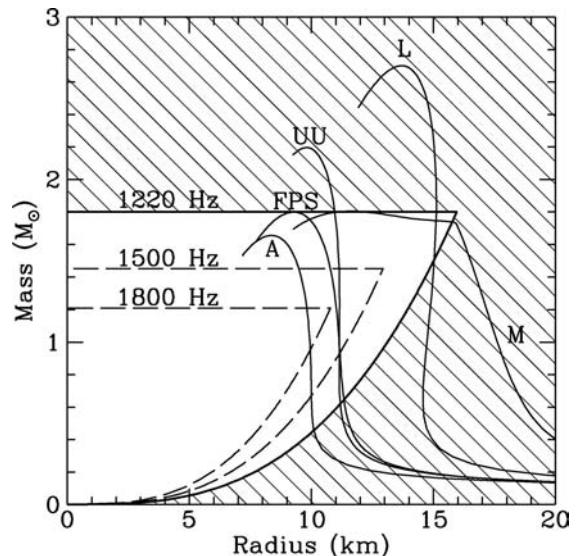


Fig 11.69 Relationship between neutron star mass and radius, according to currently acceptable models (labelled curves) for the equation of state of degenerate neutron matter. The radius is constrained by observations of kilohertz QPOs in LMXBs, the current highest being around 1330 Hz, which provides an upper limit by being at least equal to r_{ISCO} . The observations therefore constrain the neutron star’s parameters to lie in the unshaded regions (diagram from Miller et al., 1998).

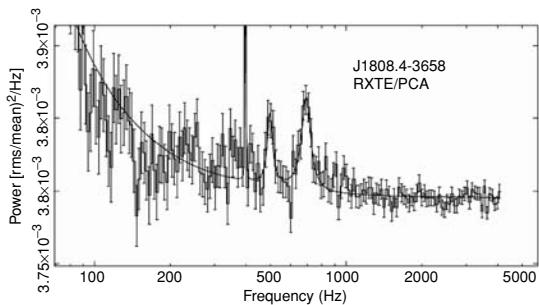


Fig 11.70 Presence of twin-peak kilohertz QPOs in this RXTE power spectrum of the millisecond X-ray pulsar, J1808.4-3658. In this case, the separation of the peaks is almost exactly half the spin frequency of the neutron star, which is 401 Hz (evident here as the sharp, extremely significant peak in the power spectrum) (diagram adapted from Wijnands et al., 2003).

11.3.6 Accreting millisecond X-ray pulsars

11.3.6.1 Neutron star signatures

Two ways of using X-rays to study neutron stars and their immediate vicinity have been discussed in some detail – the X-ray bursts that arise from nuclear reactions occurring close to the neutron star's surface in LMXBs and the regular pulsations that show the presence of a very powerful magnetic field in HMXBs. Apart from the 7-s pulsar in 4U1626-67, these two observed properties appeared by the 1990s to be cleanly divided between LMXBs (bursters) and HMXBs (pulsars). This was a puzzling situation. After all, the Rapid Burster must have a magnetosphere (which acts as the gate), and X-ray pulsations should occur as the material flows down the magnetic field lines onto the poles of the neutron star. Also, the matter flowing onto the poles in HMXB X-ray pulsars would be

Box 11.12 | Relationship to black-hole LMXBs

Similar QPOs at high frequencies have been seen from the black-hole LMXBs, which immediately indicates that, if the phenomenon has a common origin in all LMXBs, then it cannot originate on the compact object's surface. Furthermore, in the black-hole systems, they do not exhibit twin peaks either, again because the absence of an observable surface precludes the existence of the spin frequency with which to 'beat'. Without the influence of a magnetic field or direct radiation from the compact object, it might be expected that the black-hole systems would exhibit more stable high-frequency QPOs, and this is essentially what is observed. And with the r_{ISCO} being physically larger for black-hole systems (because of their larger masses), the highest frequencies observed are significantly lower than for the neutron star systems.

A further correlation of note (Fig. 11.71) allows both black-hole and neutron star LMXBs to be compared with their white dwarf cousins, the cataclysmic variables (see Chapter 10). Where both high-frequency and slow QPOs are present, their ratio is close to 15, and this extends to the ratio of dwarf nova oscillation frequencies to QPOs in CVs.

Such a relationship supports the interpretation of the QPOs as features in the inner disc which are linked to the rotating white dwarf.

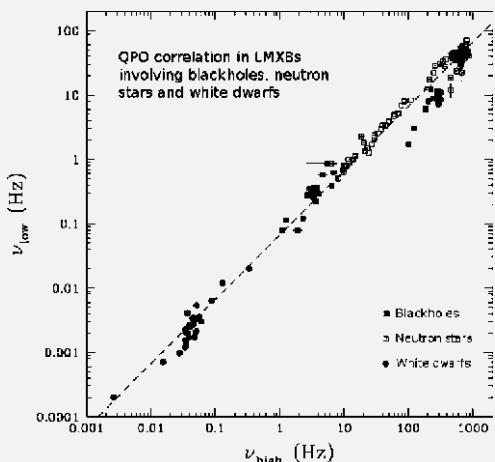


Fig 11.71 Correlations between high- and low-frequency QPOs in black-hole and neutron-star LMXBs and CVs (the latter exhibiting both QPOs and 'dwarf nova oscillations'). The ratio of these two components is close to 15 (the dashed line) (diagram adapted from Warner et al., 2003).

expected to produce occasional X-ray bursts. Why are X-ray bursts not a characteristic of all X-ray sources in which the compact object is a neutron star?

The answer has already been hinted at. The X-ray bursters do not burst *all* the time, and their bursting behaviour is related to the overall X-ray brightness level. At high mass-transfer rates, the helium-burning instability is suppressed, so bursts do not occur. Almost all of the X-ray pulsars are in highly luminous HMXBs and hence in a high mass-transfer state, in which X-ray bursts are inhibited.

11.3.6.2 Magnetic field decay

The principal difference between the pulsars and the bursters is *age*. The HMXB pulsars all have massive early-type mass donors and *must* therefore be young (because an OB star has a lifetime of only a few million years). The bursters, on the other hand, are all associated with LMXBs and are likely to be very old. Most of their mass donors are low-mass stars that could have existed in their present form for hundreds of millions, or even billions, of years. Although both HMXBs and LMXBs contain neutron stars, it is the age that is critical in determining the strength of the magnetic field. Frozen-in during the initial collapse of the star that formed the neutron star, this field gradually decays over time. The characteristics of the Rapid Burster indicate that it has a field that is weaker than in most of the pulsars. The pulsars' fields never give way to the accreting matter to allow the type II bursts to occur.

11.3.6.3 The link with millisecond radio pulsars

There is another crucial difference between HMXBs and LMXBs, and that is the mass transfer mechanism. LMXBs accrete via Roche-lobe overflow, a mechanism that transfers angular momentum into the accretion disc and then onto the compact object, causing it to spin up. The angular momentum transferred this way is *much* greater than via the stellar wind accretion in HMXBs,

where the wind is moving radially outwards. Consequently, we would expect old neutron stars in LMXBs to be spinning *extremely* rapidly, perhaps close to breakup, where the centrifugal force at the equator barely balances that of gravity. This should happen at a rotation period of about 1 ms! Remarkably, such systems have been known for almost 30 years, but they are millisecond *radio* pulsars (MSRPs). Found in the early 1980s, they have weak ($\sim 10^8$ G) magnetic fields (similar to those expected for LMXBs) and have periods as short as 1.5 ms. With no detectable supernova remnant and weak fields, these had to be old systems. The only way high spin rates could be achieved is through accretion in an LMXB. Eventually, the donor's mass drops to a very low value, the mass transfer rate falls off and the spun-up neutron star turns on as a radio pulsar. It took a while to find comparable fast X-ray pulsations in currently active LMXBs.

11.3.6.4 SAX J1808.4-3658

The situation changed dramatically with the launch of RXTE in late 1995. Within just a few months, millisecond variability had been discovered in two LMXBs, but not at strictly stable periods – they were quasiperiodic.¹ However, when the X-ray transient, SAX J1808.4-3658, became active again in 1998, the RXTE observations were spectacular (Fig. 11.72). The power spectrum revealed the very precise 401-Hz pulsation due to the spinning neutron star modulated by the 2-hour binary orbit. The properties of five of these RXTE-discovered ‘accreting millisecond X-ray pulsars’ (AMXPs) are given in Table 11.8, and there are currently 11 such systems known.

The observed properties of SAX J1808.4-3658 fit very well with those of the MSRs (Fig. 11.73, which also shows the range of magnetic fields from young and old pulsars). Furthermore, many of the AMXPs have very short orbital periods and must therefore be ‘ultra-compact X-ray binaries’ (UCXBs), in which the mass donor is itself extremely low mass (a few hundredths of a solar mass) and likely degenerate. It is easy to see how such systems will next become MSRs.

¹ There were several reasons for this. RXTE’s X-ray detectors (1) were much larger than those of EXOSAT and Ginga, (2) they could view the brightest targets directly and (3) the telemetry could allow for higher time resolution.

Table 11.8 Observed properties of accreting millisecond X-ray pulsars.

Source	P_{orb} (min)	f_{spin} (Hz)	a (light seconds)
XTE J0929-3314	43.6	185.1	0.006
XTE J1751-305	42.4	435.3	0.010
XTE J1807-294	40.1	190.6	0.005
SAX J1808.4-3658	120.9	401.0	0.063
XTE J1814-338	256.5	314.3	0.390

RXTE/PCA observations of SAX J1808.4-3658

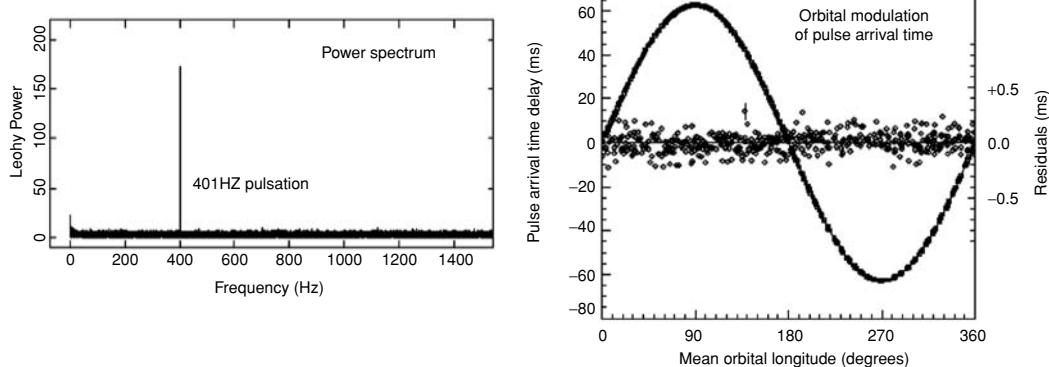


Fig 11.72 (left) Power spectrum showing the discovery of 401-Hz pulsations in the X-ray transient SAX J1808.4-3658 (from Wijnands & van der Klis, 1998), which then revealed (right) their orbital modulation on the 2-hour binary period (from Chakrabarty & Morgan, 1998).

11.3.6.5 Burst oscillations

RXTE also discovered the presence of millisecond pulsations from LMXBs during type I X-ray bursts. This result was not unexpected, as any asymmetry in the thermonuclear flash around the neutron star surface would appear as a pulsation at the spin rate. Any changes in that asymmetry would manifest themselves as an apparent drift in the pulsation rate (e.g. Fig. 11.74).

More than a dozen bursters have revealed such oscillations, and inherent in this explanation is the assumption that these periodicities are close to or actually at the neutron star spin period. However, none of these sources exhibited a pulsation during their *non-bursting* intervals. To confirm this interpretation, it was necessary to find a burst oscillation from a source that had a known, stable X-ray pulsation period. This was duly found from the original AMXP, SAX J1808.4-3658, when

it underwent a type I X-ray burst in October 2002, and an oscillation appears just after the peak at the same frequency (401 Hz) as during non-bursting X-ray emission (Fig. 11.72) (Chakrabarty *et al.*, 2003). This important result demonstrated unequivocally the validity of the rapidly spinning interpretation of almost two dozen LMXBs.

Oscillations have also been seen during *superbursts*, and if long enough, then it is possible to observe the frequency drift of the oscillation as a result of the orbital motion of the neutron star about its companion. This was detected by Strohmayer and Markwardt (2002) during an 800-s superburst from MXB1636-536.

The range of spin frequencies observed so far in AMXPs is shown in Fig. 11.75 (left), which appears to cut off near 730 Hz. There is no observational limit set by this value; indeed, RXTE/PCA could detect a 2-kHz X-ray pulsar if one existed. And it is

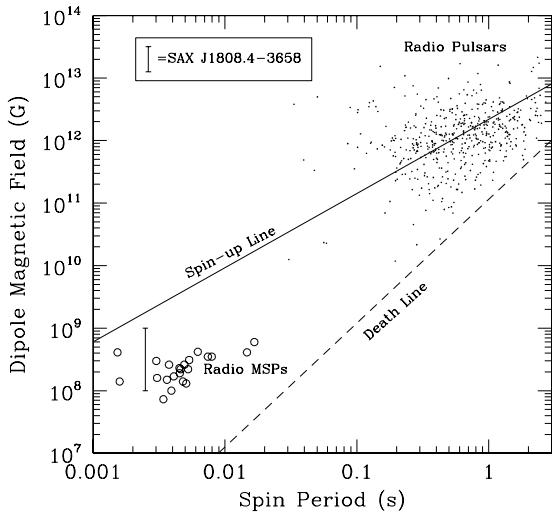


Fig 11.73 Comparison of the magnetic field and rotation rate of the accretion-driven MSP SAX J1808.4-3658 with those of slowly spinning radio pulsars (dots) and millisecond radio pulsars (open circles). Below the ‘death’ line, the pulsars would not have detectable radio emission. The spin-up line is an estimate of the period that might be expected for a given magnetic field due to the torque applied at the inner disc, where it interacts with the neutron star’s magnetosphere. See also Fig. 11.16 (diagram by Psaltis & Chakrabarty, 1999).

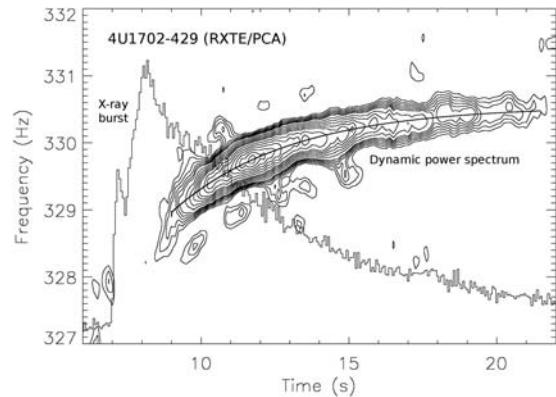


Fig 11.74 The light curve of a type I X-ray burst from 4U1702-429 observed with the RXTE/PCA. The high time resolution allowed the power spectrum to be derived as a function of time during the burst. This *dynamic power spectrum* is overplotted as a contour map. The solid curve through the contour map shows the peak of the oscillation frequency (which moves from 329 to 330 Hz), which is only detectable during the burst itself (diagram by Strohmayer & Markwardt, 2002).

unlikely that the neutron star is breaking up due to centrifugal forces at this spin rate as that would rule out essentially all currently acceptable equations of state for neutron star matter (Fig. 11.75, right). One possible explanation is that rotational energy is lost via gravitational radiation. If so,

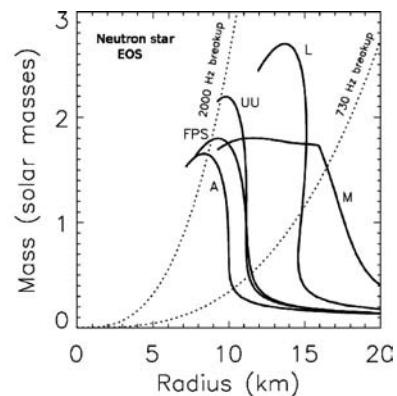
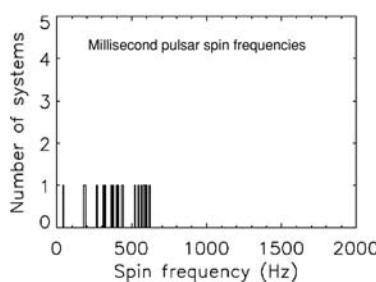
AMXPs are potential targets for gravitational wave experiments (Chakrabarty, 2008).

11.4 | Globular cluster X-ray sources

11.4.1 Background

The first X-ray all-sky surveys revealed a handful of sources that, in spite of large positional

Fig 11.75 (left) Distribution of AMXP spin frequencies, which cuts off at around 730 Hz, whereas RXTE/PCA could detect such modulations up to about 2 kHz. (right) Models for the equation of state of neutron star material with the constraints imposed by allowing neutron stars to be able to spin up to 2000 and 730 Hz. Allowed regions are to the right of the dotted lines. If the cutoff really were at 730 Hz, it would exclude virtually all currently acceptable models (diagram by Deepto Chakrabarty, MIT).



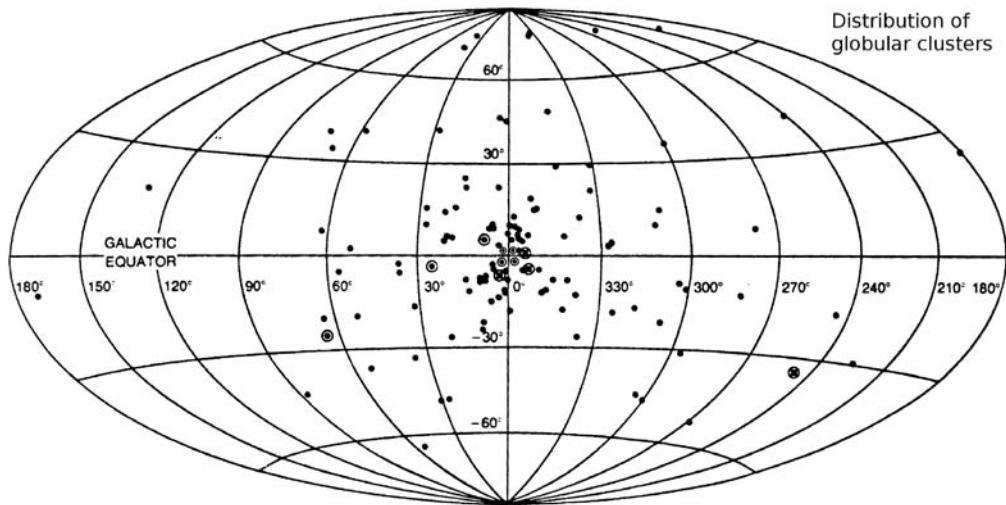


Fig 11.76 Distribution of globular clusters on the sky, in galactic co-ordinates. They are not confined to the galactic plane at all (as are most of the visible stars of our Galaxy) but instead are essentially uniformly distributed around the galactic centre. The X-ray-luminous clusters are circled (diagram adapted from an original by George Clark, MIT).

uncertainties, appeared to be associated with globular clusters (at high galactic latitudes, there were no other obvious candidates). By 1975 there were five luminous X-ray sources that had been located within globular clusters: NGC 1851, NGC 6441, NGC 6624, NGC 6712 and the well-known M15 (NGC 7078). It was a surprising result. There were a large number of X-ray sources relative to the total number of stars within the clusters. There are $\sim 10^{11}$ stars in our Galaxy, which contains only about 100 luminous X-ray sources. Maintaining that ratio for the $\sim 10^7$ stars in all the globular clusters in our Galaxy, there should be only 0.01 X-ray sources, whereas there were 5!²

Globular clusters are almost perfectly spheroidal aggregations of stars (hence the name) which are amongst the oldest objects in our Galaxy. Beautiful and impressive to observe (even through modest-sized telescopes, but quite spectacular with HST, as in Colour Plate 31) they were formed from the original proto-galactic gas cloud before the main disc of our Galaxy (and hence before the Sun) was formed.

The distribution of the 147 galactic globular clusters on the sky shows a uniform spread about

the galactic centre, with no concentration in the plane or disc (Fig. 11.76). The clusters travel in orbits about the Galaxy, which, of course, take them through the plane. During such passages, some loosely bound stars are lost, together with any cluster interstellar gas which is stripped by the higher gas density in the plane. All the stars in a given globular cluster were formed at about the same time and from the same mix of elements (mostly hydrogen, with very little in the way of heavier elements in those early days of the Galaxy). Hence a cluster's Hertzsprung-Russell diagram (Fig. 11.77) consists of a truncated main sequence together with red giant and horizontal branches containing highly evolved stars.

This well-known situation occurs because of the dependence of stellar lifetime on initial mass. The heavier the star, the shorter its 'normal' life. Those heavier than about $0.8 M_{\odot}$ have exhausted their original nuclear fuel and have evolved off the main sequence. This turn-off point depends principally on the age of the cluster, and it is clear that these clusters are indeed extremely old (typically older than our Sun, some close to the age of the Universe). More importantly, the more

² We now know, as a result of more thorough surveys, that there are 13 luminous X-ray sources associated with galactic globular clusters.

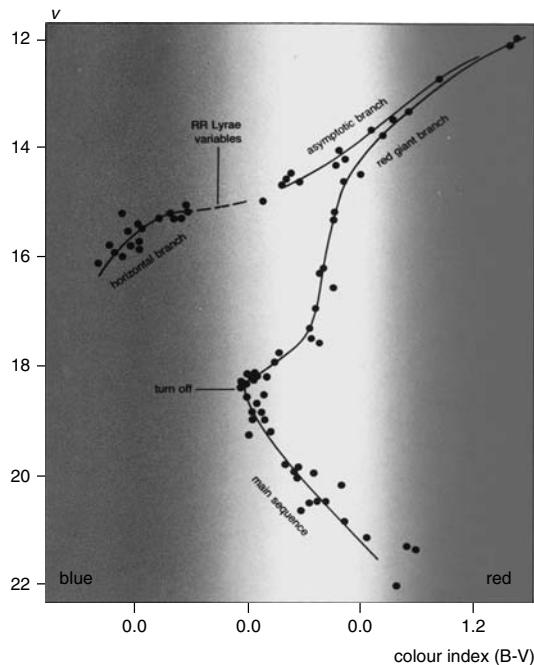


Fig 11.77 Hertzsprung-Russell diagram for the globular cluster M92, showing the turn-off from the main sequence into the red giant branch and then into the horizontal branch. The temperature is inferred from the observed colours of the cluster stars; the luminosity is simply related to the apparent brightness of the stars but requires the distance of the cluster to be known. The turn-off is in fact at an apparent magnitude of about 19 and indicates that all stars brighter than this have left the main sequence. By comparison with models of stellar evolution, the age of the cluster can be estimated (modified version of diagram in Cambridge Atlas p. 291).

massive original members of the cluster, after supernova explosions, will have left behind stellar remnants in the form of neutron stars and black holes. Thus the clusters contain just the objects needed to form X-ray sources.³

Globular clusters continue to demonstrate their importance to models of stellar evolution, with the frequent occurrence of X-ray binaries and millisecond radio pulsars. Clearly the process by which X-ray sources are formed in globular clusters is greatly enhanced compared to the rest of

our Galaxy. To understand why, we need more details about the nature of the X-ray sources themselves, and the observed properties of the luminous sources are summarised in Table 11.9. That they are LMXBs is an inevitable consequence of the age of the clusters. And the discovery of X-ray bursts from virtually all the luminous globular cluster X-ray sources demonstrated that they must contain neutron stars.

11.4.2 X-ray bursts and precise X-ray locations

The first X-ray burster was discovered in the globular cluster NGC 6624, and similar behaviour was quickly found in several of the other cluster sources. This identification constrains the mass of the compact object to be close to $1.4 M_{\odot}$. All the strong globular cluster sources display bursting behaviour, except one, X-1 in M15, which is an accretion disc corona source, and so any bursts from this source might well be washed out.

The mass range was further constrained when the Einstein Observatory HRI obtained accurate (<3 arcsec) X-ray positions showing that *all* the globular cluster X-ray sources were located very close to the cluster centre (Fig 11.78) (Grindlay *et al.*, 1984). Indeed, this work, subsequently enhanced considerably by Chandra observations, has showed that by far, the majority of all X-ray sources were within 2 core radii of the cluster centre (Fig. 11.79).

This could not happen by chance. The X-ray objects had to be different from the other stars in the cluster. They were heavier, and mass segregation has led to them settling closer into the core. A detailed statistical analysis showed that the X-ray sources were between 2 and 3 times as heavy as the average cluster star, which is about $0.5 M_{\odot}$. Thus the cluster X-ray sources have an average mass of about $1.5 M_{\odot}$, very close to the canonical mass of a neutron star, exactly as expected given that they are all X-ray bursters. It also implies

³ The reader who is familiar with the physics of the formation of neutron stars in supernova explosions may challenge this assertion on the grounds that radio pulsars are known to have quite large space motions, acquired during the supernova explosion itself, which are typically much larger than the escape velocity from a globular cluster. Whilst this is true, there are enough pulsars observed with low space velocities to suggest that a fraction of new neutron stars formed (believed to be between 10% and 20%) will be retained by the parent globular cluster.

Table 11.9 Properties of luminous X-ray sources in globular clusters.

Source	Cluster	$\log L_X$ (ergs s $^{-1}$)	Distance ^a (kpc)	P_{orb} (hours)	Notes
X0513-401	NGC 1851	36.1	12.1	0.28	UCB
X1724-308	Terzan 2	36.7	8.7		
MXB1730-335	Liller 1	36.8	9.6		Rapid Burster, recurrent transient
X1732-304	Terzan 1	36.8	5.6		Transient
X1745-25	Terzan 5	35.3	7.6		Transient
MX1746-20	NGC 6440	37.0	8.4		Transient (outbursts 1971, 1998, 2001)
4U1746-37	NGC 6441	36.8	11.7	5.7	
X1751-31	Terzan 6	36.8	9.5	12.4	Transient, eclipsing
4U1820-30	NGC 6624	38.0	7.6	0.19	UCB, possible triple
X1836-33	NGC 6652	36.0	10.1		
X1850-086	NGC 6712	36.4	6.9	0.34	UCB
4U2127+11 X-1	M15	36.3	10.3	17.1	ADC (optical counterpart AC211)
4U2127+11 X-2	M15	36.1	"	0.38	UCB

Note. UCB = ultra-compact binary.

^a See Harris (1996).

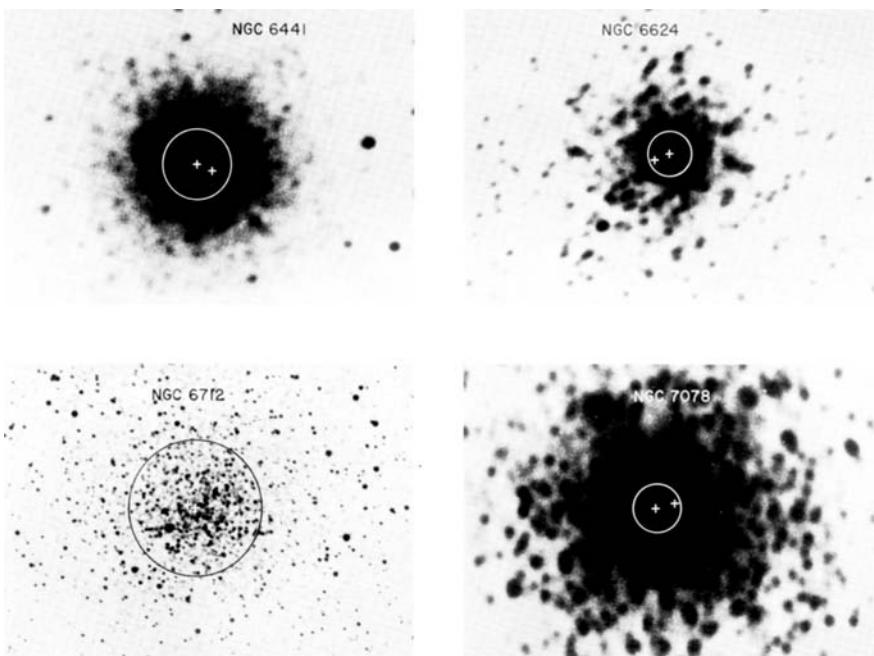


Fig 11.78 Montage of optical photos of centres of some bright globular cluster X-ray sources. The circle shows the core of the cluster as defined by the radius at which the brightness has fallen to half its peak value. The cross at the centre of the circle therefore represents the best estimate of the optical centre of the cluster. The Einstein HRI X-ray position is the offset cross, and its uncertainty is the size of the cross itself (Images by Josh Grindlay, CfA).

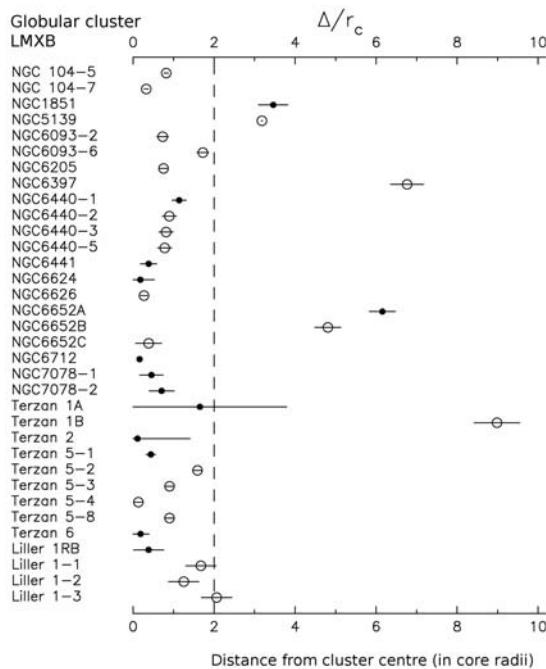


Fig 11.79 A compendium of all galactic LMXB locations within globular clusters, expressed as their distance from the centre of the cluster in units of the core radius, r_c . The luminous X-ray sources are plotted as filled circles, faint sources with open circles (adapted from Verbunt & Lewin, 2006).

that they have very low mass companion stars.

11.4.3 Formation of binary X-ray sources in clusters

Having established the mass through their locations, the globular-cluster X-ray sources were identified as LMXBs. Nothing else was capable of delivering the observed X-ray power, and in any case, they were otherwise very similar (X-ray spectrum, X-ray bursts, distribution about the galactic centre) to the X-ray sources of the Galactic Bulge. However, the environments of globular clusters are clearly highly conducive to compact binary formation. This must be related to the extremely high stellar densities in the cluster cores, and there are two mechanisms that have been examined in detail (see Fig. 11.80) by which a lone neutron star

can acquire a companion star in such an environment:

1. *Compact object captures single star.* Unless the two stars physically collide (which, though unlikely, can happen), they have too much kinetic (motion) energy to be trapped into a stable binary orbit. However, as pointed out by Fabian *et al.* (1975), there is a mechanism by which some of this excess kinetic energy can be dissipated, and that is *tides*. As the compact object approaches the normal star, it distorts it substantially (raises tides), which absorbs energy (in exactly the same way that our Moon raises tides in the Earth's oceans). If this happens quickly enough, then the two stars become bound.

2. *Compact object interacts with a wide binary, a 'triple-star' event.* A wide (non-interacting) binary of two normal stars is approached by the compact object, and a complex triple-star interaction takes place. This results in the compact object replacing the lower mass star in the binary and ejecting it.⁴

Although the cross-section (likelihood of encounter) is much greater for the triple-star event (the compact object only has to hit somewhere between the two normal stars), the problem is that such wide binaries appear to be extremely rare in globular clusters. Calculations show, therefore, that process (1), the tidal capture mechanism, is expected to be by far the more dominant (Verbunt & Hut, 1987). It is also clear in that case that the capture is more likely to happen if the normal star is physically larger. This means that an evolving star (onto the red giant branch) has a greater chance of being captured.

It is now clear from detailed numerical modelling that binary and triple encounters inside globular cluster cores are extremely important in controlling the evolution of the cluster. It was once thought inevitable that the high central density would lead to collapse of the cluster and formation of a massive black hole. Instead, the formation of binaries by the processes just described can halt, and even reverse, the collapse. Stars ejected in triple-star interactions can receive sufficient 'kick' to completely leave the cluster. It has even been suggested that all LMXBs are formed inside

⁴ This is akin to the exchange collisions that occur in nuclear particle reactions.

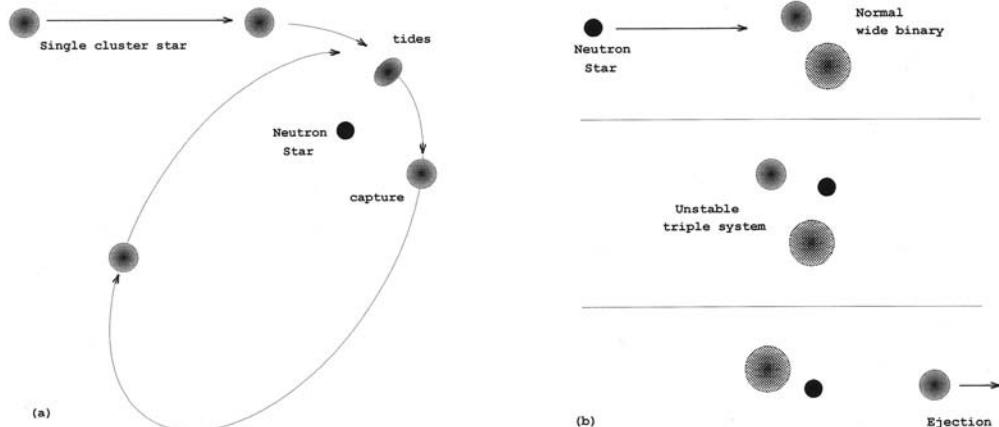


Fig 11.80 Schematic of the processes of binary formation involving compact objects in globular clusters. (left) In the tidal interaction of two stars, some of the relative kinetic energy must be lost for them to become bound. This is done through the tides that are raised on the normal star by the compact object. Since the initial orbit in such a capture is eccentric, these tides continue to operate at each close approach, thereby causing the orbit to rapidly circularise. (right) The alternative mechanism involves a three-body interaction in which the compact object joins an existing wide binary of two normal stars. This is much more complex. Eventually, one of the normal stars in this temporary triple system is ejected, thereby leaving the compact object with the other as a companion (based on original diagrams by George Clark, MIT).

globular clusters, which subsequently evaporate, but this is an area of active current research, and there is far from agreement on the issue.

Demonstrating conclusively that cluster X-ray sources are LMXBs proved challenging, with only two periodicities being uncovered in the first 25 years of globular cluster X-ray source studies. It is worth looking at these two more closely as both are remarkable when compared with the LMXB population as a whole.

11.4.3.1 NGC 6624 (4U1820-30): The shortest orbital period known

As well as being the brightest cluster X-ray source and the first burster, 4U1820-30 was also the first of this class to reveal an X-ray periodicity. As part of a study of X-ray QPOs in globular cluster sources, Stella *et al.* (1987) obtained extensive observations of 4U1820-30 with EXOSAT (Fig. 11.81). To their surprise, they found a weak (few per cent amplitude) but precise modulation at a period of 11.4 min.

At first sight, such a discovery does not appear to be remarkable. Neutron stars with strong magnetic fields give rise to X-ray pulsations as

they rotate, with periods in the range of seconds to minutes. However, one important characteristic of X-ray pulsars is that the rotation period changes with the luminosity of the source (more accreting material spins the neutron star up faster, whereas less material spins it down; see Fig. 11.16). In the case of 4U1820-30 the periodicity was found to be present in earlier observations (some going back to the mid 1970s), with exactly the same value, even though this object is extremely bright in X-rays and must therefore be accreting at a very high rate.

The implication is that the 685-s period must be orbital in origin as no other explanation appears viable. 4U1820-30 has the shortest orbital period known and must be small in size – so small, in fact, that no normal star could fit within it as the mass donor. Only a degenerate object can fit in the space available, and the current model consists of a $0.05 M_{\odot}$ helium white dwarf (which is only $0.03 R_{\odot}$ in size) filling its Roche lobe and transferring material onto a $1.5 M_{\odot}$ neutron star. It is an *ultra-compact binary*, or UCB. 4U1820-30 would fit entirely inside our Sun, with room to spare!

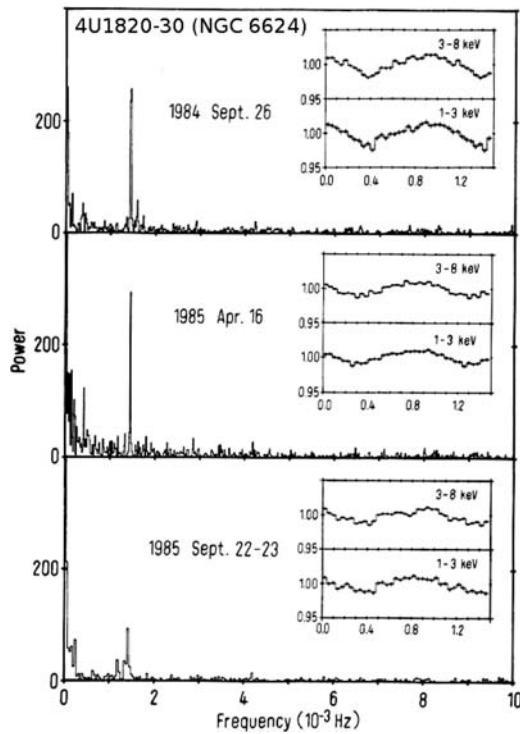


Fig 11.81 The shortest orbital period known. The EXOSAT discovery of this 685 s period of 4U1820-30 (NGC 6624) is revealed in these power spectra of the 1–3 keV X-ray data from three observations in 1984 and 1985. The folded light curve for each observation shows the low amplitude of this modulation, normally less than 3 per cent, which is why it was not discovered by earlier observations. The fact that the modulation does not change frequency between observations indicates that the periodicity is orbital and not simply due to the rotation of the neutron star (diagram from Stella *et al.*, 1987).

11.4.3.2 M15 (4U2127+11): Not one LMXB, but two!

The massive, centrally condensed globular cluster M15 provides an ideal environment for the formation of exotic binaries. It contains a luminous LMXB (4U2127+11), which was the first to be optically identified in a globular cluster, and six millisecond-pulsars in the core. Its surprisingly bright (at 15th magnitude) optical counterpart is the 17.1-hour eclipsing binary, AC211, which is by far the most optically luminous of *any* galactic LMXB. Its broad eclipse feature (at all wavelengths) suggests a high-inclination, accretion disc corona (ADC) system.

This would account for the high optical luminosity associated with a relatively faint X-ray source, as the latter would only be scattered X-rays from the corona. However, a long-standing puzzle had been the detection of extremely luminous type I X-ray bursts from M15, as this would imply that the compact object was directly visible.

The problem of the ‘mixed’ X-ray and optical properties of the M15 luminous X-ray source was elegantly solved by Chandra and HST. There is not one LMXB in the core of M15, but two! Figure 11.82 shows how Chandra’s exquisite spatial resolution revealed that the two LMXBs are a mere 2.7 arcsec apart (White & Angelini, 2001; Hannikainen *et al.*, 2005), and HST’s far-UV imaging locates the counterparts to both sources. Furthermore, Dieball *et al.* (2005) used HST to show that X-2 was also a UCB, with a 22.6-min period.

11.4.4 Low-luminosity globular cluster X-ray sources

The first faint X-ray sources ($<10^{35}$ ergs s^{-1}) were detected by the Einstein Observatory, and whilst a few more were added by ROSAT, the numbers have increased dramatically with Chandra and XMM observations (Fig. 11.83). These fainter sources encompass a much wider range of X-ray-emitting systems (Fig. 11.84).

The number of such low- L_X systems in globular clusters observed by XMM and Chandra is literally in the hundreds (see e.g. Pooley, 2006). A superb example is Chandra’s image of 47 Tuc (Colour Plate 32), a relatively nearby (5-kpc distant) cluster that contains 22 MSPs along with all the other types summarised in Fig 11.85. With such deep imaging, though (to reach the faintest members of the cluster), comes the possibility of picking up a mix of non-cluster sources through foreground stellar coronae and background quasars, especially in the less-dense outer regions of a cluster.

11.4.5 Origin and evolution of globular cluster X-ray sources

M15 and NGC 6624 thus contain three LMXBs, two of which are UCBs. With additional help for studying UV/optical counterparts from HST’s spatial resolution, and the fact that UV-bright accretion

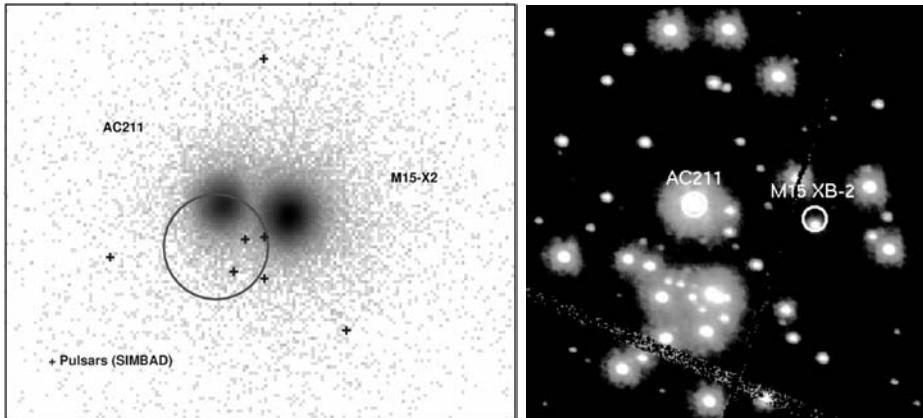


Fig 11.82 (left) Chandra HRC image of M15 shows clearly the presence of two luminous LMXBs, a mere 2.7 arcsec apart. M15's highly concentrated core, which has a radius of 2.2 arcsec, is represented by the circle. The locations of seven (of the eight known) millisecond radio pulsars in M15 are marked by crosses. (right) HST FUV image of the same region shows how AC211 (X-1) is the dominant emitter in the UV from the M15 core. It also identifies the faint counterpart of X-2 (images from Hannikainen et al., 2005; Dieball et al., 2005).

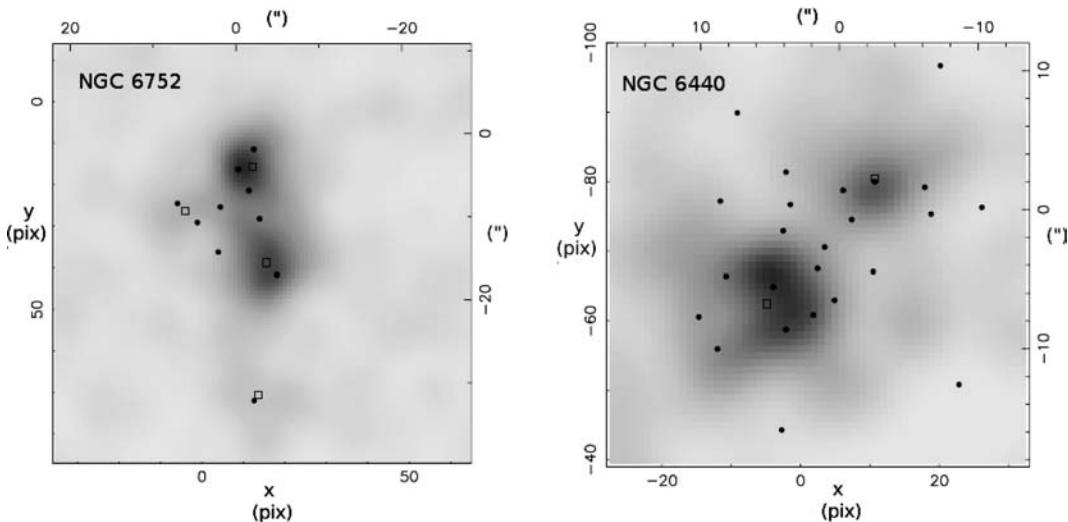


Fig 11.83 Comparison of the imaging performance of ROSAT and Chandra in the crowded regions of globular clusters (left) NGC 6752 and (right) NGC 6440. The ROSAT images are shown in grey scale, with the detected point sources indicated with open squares; the filled circles show those detected with Chandra. Clearly many of the 'single' ROSAT sources have been found by Chandra to be multiple (diagram adapted from Verbunt, 2005).

discs stand out against the surrounding cool, red giants, the number of cluster LMXBs with known orbital periods has grown to seven (Table 11.9 and Fig. 11.85). Remarkably, four of these are UCBs, an extremely high fraction compared to the galactic population of LMXBs as a whole, with 4U1820-30's 11-min binary being the shortest and most exotic of them all. Indeed, there are less than a dozen

such UCBs known in total. What is the evolutionary path by which such systems form, and why is it so enhanced in globular clusters?

The M15 and 47 Tuc images (Fig. 11.82 and Colour Plate 32) demonstrate the presence of a large pool of compact objects circulating within the dense cores of these clusters. Obviously, these are remnants of an earlier population of massive

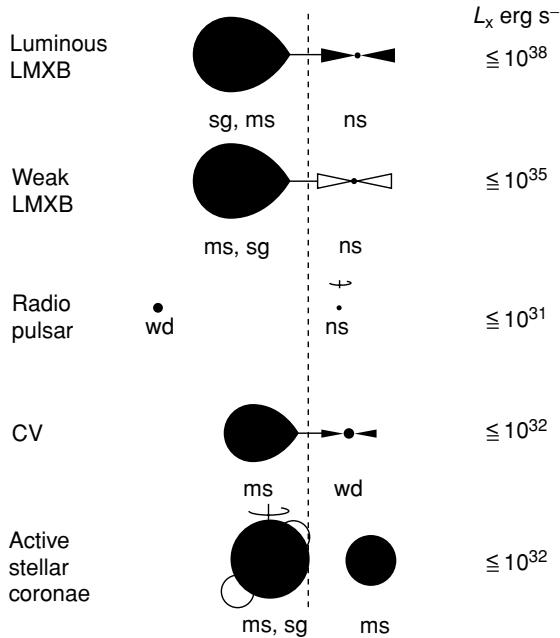


Fig 11.84 Types of X-ray sources that have been found in globular clusters. These range from the luminous LMXBs that are the main topic of this section through LMXBs in a quiescent state to binary radio pulsars, CVs and active stellar coronae. Typical maximum luminosities are indicated along the right (diagram adapted from Verbunt & Lewin, 2006).

stars within the cluster, and they indicate that while some are expected to have been ejected from the cluster in their supernova explosions, many remain. There is also, as yet, no evidence for the presence of an ‘intermediate-mass’ ($\geq 1000 M_{\odot}$) black hole at the centre of M15.

These roaming neutron stars will most likely interact with main sequence stars (Fig. 11.86, top). In one-third of the collisions, the neutron star actually scores a direct hit on the star, and numerical simulations indicate that the normal star will be totally disrupted, forming a massive disc around the compact object. Eventually, this disc will be accreted (or expelled by radiation pressure), leaving behind the lone neutron star. In the other two-thirds of the interactions, the star is captured by the neutron star, forming an LMXB with an orbital period of several hours, such as X-1/AC211 in M15.

But what happens when a neutron star meets a red giant? As outlined by Verbunt (1987), red giants are rarer but are more likely to be found in the core and are much larger stars (and hence better targets for compact objects to hit). About 10 per cent of collisions are thought to involve red giants. A direct hit (again expected in one-third of the events) does not disrupt the star because the core is small and dense. Box 11.13 explains how a

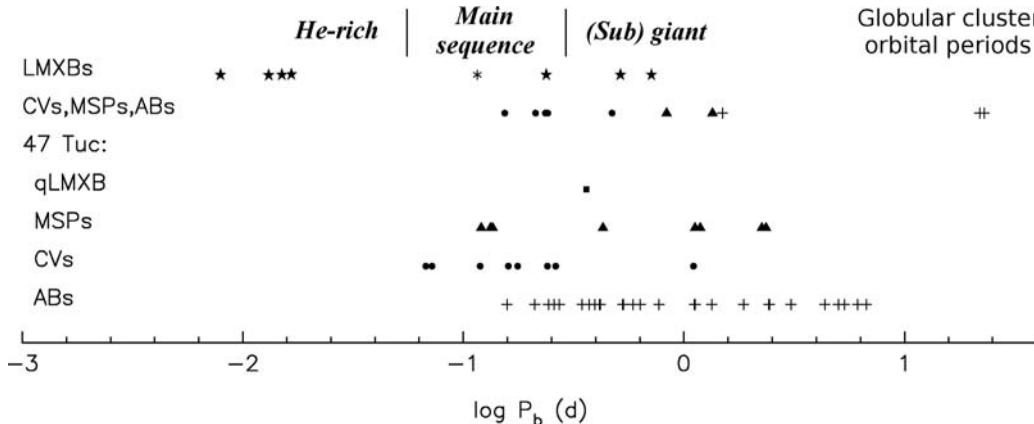


Fig 11.85 Distribution of orbital periods for binaries in galactic globular clusters, with the top line representing the luminous LMXBs (including the one period determined for an M31 cluster source, marked with an asterisk). The remaining lines are the low-luminosity X-ray sources of various types (CVs, cataclysmic variables; MSPs, millisecond pulsars; ABs, active binaries; qLMXB, quiescent LMXBs), with 47 Tuc shown separately since it has the most number of periods determined of any cluster. The nature of the mass donor is a function of the orbital period and is indicated (with approximate divisions) at the top (diagram adapted from Verbunt & Lewin, 2006).

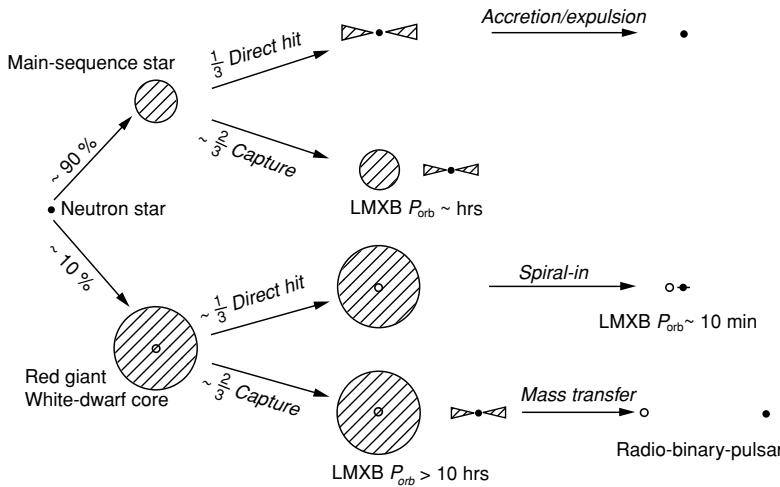


Fig 11.86 Evolutionary scenarios for different ways of capturing a neutron star in the crowded core regions of globular clusters. Most encounters involve main sequence stars, the majority of which will form typical LMXBs. Alternatively, a very short period LMXB can be formed by a direct hit with a red giant star in which the neutron star spirals in through the red giant's extended atmosphere. The end results of all these scenarios have already been observed (based on a diagram by Frank Verbunt).

Box 11.13 The structure of a red giant

A red giant forms when a normal star has consumed a significant fraction of its main nuclear fuel (hydrogen). The central helium core, no longer supported by internal nuclear energy, then begins to contract under its own weight. Hydrogen burning continues, however, in a shell around the core.

As the core contracts, the surface gravity in the shell increases. This in turn increases the rate of energy generation in the shell. The increased luminosity causes the outer parts of the star to expand and become less dense. Because the radiation we see from the star is being emitted from a larger area (than a normal star), the surface appears to be cooler or redder; hence the name 'red giant'. In fact, the centre of a red giant is much hotter and denser than a normal star. The star is becoming a double entity: a compact central core (the progenitor of a white dwarf) surrounded by a large, extremely diffuse envelope.

red giant is physically very different from a normal star.

Finding itself in the relatively dense outer layers of the red giant, the neutron star slows and spirals in towards the compact centre of the red giant (which will be between 0.1 and $0.5 M_{\odot}$ in size and composed almost entirely of helium, the

nuclear 'ash' from the red giant's earlier life). During this (short-lived) process, the neutron star accretes material, releasing copious X-rays that disperse the red giant's atmosphere. It is possible that M15/AC211 is in exactly this phase now. At the end of the spiral-in phase, the neutron star and helium white dwarf would be left in a tight orbit with a period of about 10 min. This would correspond to the present state of 4U1820-30.

Alternatively, if the neutron star captures a main sequence star that is just at the turn-off point in its evolution (Bailyn & Grindlay, 1987) or a sub-giant (Podsiadlowski *et al.*, 2002), then the subsequent mass transfer will be unstable, i.e. the loss of a small amount of material from the companion will lead to the star further overfilling its Roche lobe, hence transferring more mass, and so on, and the outer envelope of the star will rapidly expand to encompass the neutron star, as well. At this point, which is again suggested to be the M15/AC211 case, the situation is almost indistinguishable from that postulated for the neutron star-red giant collision, and the subsequent evolution is the same.

11.4.5.1 The final phase: Millisecond radio pulsars

The story related to these extreme scenarios is not yet over. When the mass transfer phase is completed, a lone, or at least detached, binary

neutron star will be left. This neutron star, created in the early days of the globular cluster from a massive star, was already old. Its magnetic field would have decayed and would therefore have had relatively little effect on the accretion process after the capture of another star. The accretion disc that was formed thus extended down virtually to the surface of the neutron star, and eventually the transfer of angular momentum from the accreted material (requiring about $0.1 M_{\odot}$ of matter) would have spun up the neutron star to the speed of material in the innermost orbit of the accretion disc, which is just a few milliseconds.

At such a period the neutron star is re-born (or ‘re-cycled’) as a radio pulsar, and beginning in the 1980s, searches for very rapid radio pulsars in globular clusters have been very successful (see e.g. Manchester, 2006). These include M15 (Fig. 11.82) and 47 Tuc (Colour Plate 32). All these pulsars (a mix of single and binary sources) have rotation periods shorter than 10 ms, down to a minimum of 1.6 ms (a record held by two systems, PSR1957+20 and PSR1937+21). The re-born pulsar (such as PSR1957+20) is emitting a powerful pulsar wind that calculations suggest will ablate its companion star, leading to its complete evaporation within 10 million years. This process generates shocks near the (degenerate) companion star that produce (weak) X-rays, leading to their detection as low-luminosity sources in clusters (Fig. 11.87). It is likely that PSR1937+21, the other very fast but single pulsar, has already gone through this phase and evaporated its companion. All such objects must have gone through an LMXB stage at some point (Fig. 11.86) and have been spun up by the mass lost as the companion star evolved.

Finally, the globular cluster environment also facilitates the formation of binary radio pulsars, systems consisting of two neutron stars. There is now evidence that the LMXBs and MSPs ‘swap’ partners via exchange interactions with lone neutron stars in the dense environments of cluster cores (Grindlay *et al.*, 2006). These interactions (Fig. 11.88) lead to scatterings from the core, which is the only way to explain the properties of certain double-pulsar systems. Interestingly, such double-NS systems are expected to evolve to shorter periods via gravitational wave emission, eventually

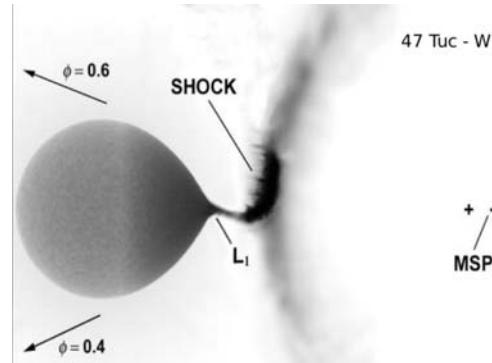


Fig 11.87 Schematic of the 47 Tuc MSP observed by Chandra (see Colour Plate 32). The mass lost from the donor is immediately shocked by the high-energy pulsar wind emanating from the MSP, producing hard X-rays. The cross marks the centre of mass, which is very close to the neutron star. This system is very similar to SAX J1808.4-3658 (based on an original diagram by Josh Grindlay).

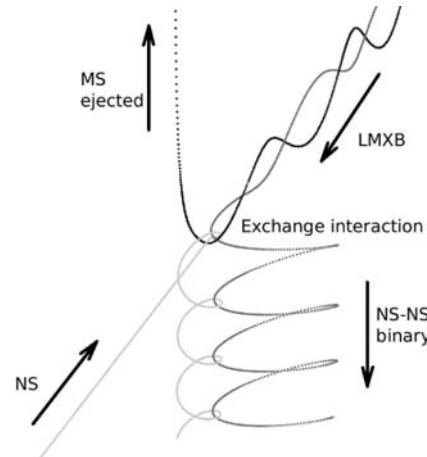


Fig 11.88 In the dense environment of a globular cluster core, an LMXB (upper right) encounters a lone neutron star (NS, lower left) and undergoes an exchange interaction. This results in the main sequence star (MS) from the LMXB being ejected and the formation of an eccentric NS-NS binary (based on an original diagram by Josh Grindlay).

coalescing and creating short γ -ray bursts (Chapter 17).

Hence the globular-cluster LMXBs and the millisecond radio pulsars are part of a continuum on the binary evolutionary sequence (see the latter stages of Fig. 11.88), providing important clues on how globular clusters form and remain stable

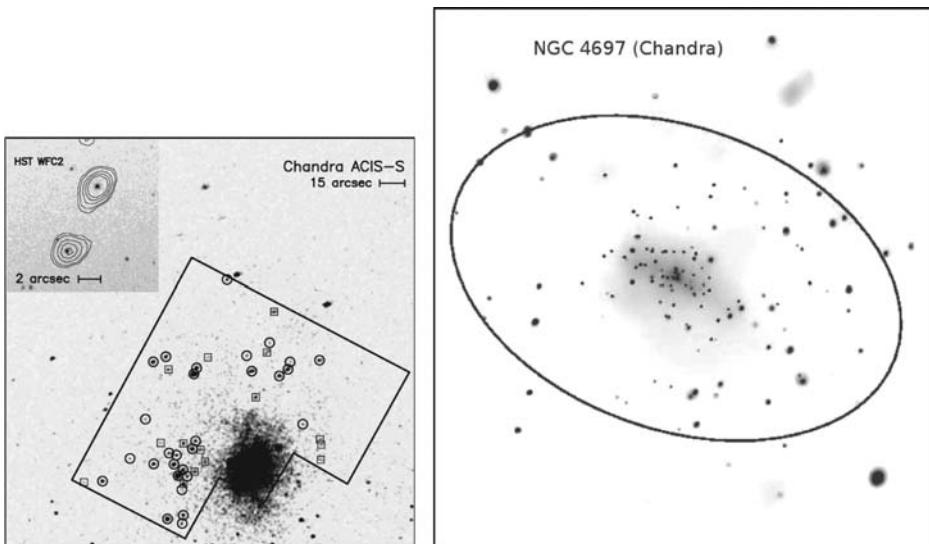


Fig 11.89 Chandra images of (left) NGC 1399 in Fornax, on which is superposed an outline of the HST/WFPC2 pointing. The circled X-ray sources were found to be associated with this giant elliptical's globular clusters (an example of which is shown in the inset), (right) NGC 4697, on which the ellipse representing the extent of the galaxy has been drawn. Many of the individual sources are associated with globular clusters (images by Angelini et al., 2001; Sarazin et al., 2001).

over enormous timescales. Their rich population of stellar exotica continues to inspire advances in our understanding of compact binary stellar evolution.

11.4.6 Globular cluster LMXBs in external galaxies

Given how difficult it has been to make progress with LMXBs in globular clusters in our own Galaxy, studying those in external galaxies is only possible with Chandra, as the highest possible spatial resolution is required. However, the combination of their faintness (requiring long exposures) and Chandra's sensitivity means that significant fractions (reaching $\sim 50\%$) of the sources detected are foreground or background objects. Hence the only way of making progress in such studies is via a program of identification of the galaxy's globular clusters with the individual X-ray sources (Fig. 11.89).

With these and many other observations of external galaxies, both spiral and elliptical, come the following key results:

- Elliptical galaxies (relative to their mass) have twice the number of globulars as spiral galaxies.

- Most X-ray sources in ellipticals are LMXBs (as there are no high-mass stars).
- Some of the globular cluster sources have a high-enough L_X that they may be BH systems.
- LMXBs are found preferentially in bright globulars, indicating that they scale with the mass of the cluster (which then produces more collisions).
- In spirals, most X-ray sources are *not* found in clusters, whereas in ellipticals, there are roughly equal numbers in clusters and the disc of the Galaxy. The latter may then be consistent with all being formed inside globulars.

References

- Abell, G. O. & B. Margon (1979) A kinematic model for SS433, *Nature* **279**, 701.
 Angelini, L., M. Loewenstein & R. Mushotzky (2001) The X-Ray globular cluster population in NGC 1399, *ApJ* **557**, L35.
 Bailyn, C. D. & J. E. Grindlay (1987) On the evolution of tidal capture X-ray binaries – 4U 2127+12 (M15) to 4U 1820-30 (NGC 6624) *ApJ* **316**, L25.

- Belian, R. D., J. P. Conner & W. D. Evans (1976) The discovery of X-ray bursts from a region in the constellation Norma, *ApJ* **206**, L135.
- Bildsten, L., et al. (1997) Observations of Accreting pulsars, *ApJS* **113**, 367.
- Blundell, K. M., et al. (2001) Images of an equatorial outflow in SS 433, *ApJ* **bf 562**, L79.
- Boirin, L., et al. (2005) A highly-ionised absorber in the X-ray binary 4U 1323-62: A new explanation for the dipping phenomenon, *A&A* **436**, 195.
- Bondi, H. & F. Hoyle (1944) On the mechanism of accretion by stars, *MNRAS* **104**, 273.
- Chakrabarty, D. (2008) The spin distribution of millisecond X-ray pulsars, *AIPC* **1068**, 67.
- Chakrabarty, D. & E. H. Morgan (1998) The two-hour orbit of a binary millisecond X-ray pulsar, *Nature* **394**, 346.
- Chakrabarty, D., et al. (2003) Nuclear-powered millisecond pulsars and the maximum spin frequency of neutron stars, *Nature* **424**, 42.
- Clark, D. H. & P. G. Murdin (1978) An unusual emission-line star/X-ray source/radio star, possibly associated with an SNR, *Nature* **276**, 44.
- Corbet, R. H. D. (1984) Be/neutron star binaries – A relationship between orbital period and neutron star spin period, *A&A* **141**, 91.
- Cornelisse, R., et al. (2000) The longest thermonuclear X-ray burst ever observed? A BeppoSAX Wide Field Camera observation of 4U 1735-44, *A&A* **357**, L21.
- Cornelisse, R., et al. (2002) A four-hours long burst from Serpens X-1, *A&A* **382**, 174.
- Cottam, J., F. Paerels & M. Mendez (2002) Gravitationally redshifted absorption lines in the X-ray burst spectra of a neutron star, *Nature* **420**, 51.
- Davidson, K. & J. P. Ostriker (1973) Neutron-star accretion in a stellar wind: Model for a pulsed X-ray source, *ApJ* **179**, 585.
- Dieball, A., et al. (2005) An ultracompact X-ray binary in the globular cluster M15 (NGC 7078), *ApJ* **634**, L105.
- Fabian, A. C. & M. J. Rees (1979) SS 433 – A double jet in action, *MNRAS* **187**, 13P.
- Fabian, A. C., J. E. Pringle & M. J. Rees (1975) Tidal capture formation of binary systems and X-ray sources in globular clusters, *MNRAS* **172**, 15P.
- Galloway, D. K., et al. (2003) Eddington-limited X-ray bursts as distance indicators. I. Systematic trends and spherical symmetry in bursts from 4U 1728-34, *ApJ* **590**, 999.
- Galloway, D. K., et al. (2004) Periodic thermonuclear X-ray bursts from GS 1826-24 and the fuel composition as a function of accretion rate, *ApJ* **601**, 466.
- Giacconi, R., et al. (1971) Discovery of periodic X-ray pulsations in Centaurus X-3 from Uhuru, *ApJ* **167**, L67.
- Grimm, H.-J., M. Gilfanov & R. Sunyaev (2002) The Milky Way in X-rays for an outside observer. Log(N)-Log(S) and luminosity function of X-ray binaries from RXTE/ASM data, *A&A* **391**, 923.
- Grimm, H.-J., M. Gilfanov & R. Sunyaev (2003) High-mass X-ray binaries as a star formation rate indicator in distant galaxies, *MNRAS* **339**, 793.
- Grindlay, J., et al. (1976) Discovery of intense X-ray bursts from the globular cluster NGC 6624, *ApJ* **205**, L127.
- Grindlay, J. E., et al. (1984) Determination of the mass of globular cluster X-ray sources, *ApJ* **282**, L13.
- Grindlay, J., S. Portegies Zwart & S. McMillan (2006) Short gamma-ray bursts from binary neutron star mergers in globular clusters *Nat Phys* **2**, 116.
- Hannikainen, D. C., et al. (2005) The X-ray source population of the globular cluster M15: Chandra high-resolution imaging, *MNRAS* **357**, 325.
- Harris, W. E. (1996) A catalog of parameters for globular clusters in the Milky Way, *AJ* **112**, 1487.
- Jimenez-Garate, M. A., J. C. Raymond & D. A. Liedahl (2002) The structure and X-ray recombination emission of a centrally illuminated accretion disk atmosphere and corona, *ApJ* **581**, 1297.
- Kawashima, K. & S. Kitomoto (1996) Photoionised plasma in Cygnus X-3, *PASJ* **48**, L113.
- Lamb, F. K., et al. (1985) Quasi-periodic oscillations in bright galactic-bulge X-ray sources, *Nature* **317**, 681.
- Lattimer, J. M. & M. Prakash (2005) Ultimate Energy density of observable cold baryonic matter, *Phys Rev Lett* **94**, 111101.
- Liedahl, D. A. & F. Paerels (1996) Photoionization-driven X-ray line emission in cygnus X-3, *ApJ* **468**, L33.
- Liu, Q. Z., J. van Paradijs & E. P. J. van den Heuvel (2006) Catalogue of high-mass X-ray binaries in the Galaxy (4th edition), *A&A* **455**, 1165.
- Liu, Q. Z., J. van Paradijs & E. P. J. van den Heuvel (2007) A catalogue of low-mass X-ray binaries in the Galaxy, LMC, and SMC (fourth edition), *A&A* **469**, 807.
- Manchester, R. N. (2006) Binary and millisecond radio pulsars, *AdSpR* **38**, 2709.
- Maraschi, L., A. Treves & E. P. J. van den Heuvel (1976) B-emission stars and X-ray sources, *Nature* **259**, 292.
- Margon, B. (1984) Observations of SS 433, *ARA&A* **22**, 507.
- Marshall, H., C. R. Canizares & N. S. Schulz (2002) The high-resolution X-ray spectrum of SS 433 using the Chandra HETGS, *ApJ* **564**, 941.
- Milgrom, M. (1978) On the nature of the galactic bulge X-ray sources, *A&A* **67**, L25.
- Milgrom, M. (1979) SS433 – The acceleration and collimation mechanisms, *A&A* **78**, L9.

- Miller, M. C., F. K. Lamb & D. Psaltis (1998) Sonic-point model of kilohertz quasi-periodic brightness oscillations in low-mass X-ray binaries, *ApJ* **508**, 791.
- Paczynski, B. (1971) Evolutionary processes in close binary systems, *ARAA* **9**, 183.
- Podsiadlowski, P., S. Rappaport & E. Pfahl (2002) Evolutionary sequences for low- and intermediate-mass X-ray binaries, *ApJ* **565**, 1107.
- Pooley, D. (2006) X-ray binary systems in globular clusters, *ESA SP* **604**, 125.
- Psaltis, D. & D. Chakrabarty (1999) The disk-magnetosphere interaction in the accretion-powered millisecond pulsar SAXJ1808.4-3658, *ApJ* **521**, 332.
- Sako, M., et al. (1999) The X-ray spectrum and global structure of the stellar wind in vela X-1, *ApJ* **525**, 921.
- Sako, M., et al. (2003) Structure and dynamics of stellar winds in high-mass X-ray binaries, arXiv:astro-ph/0309503.
- Sarazin, C. L., J. A. Irwin & J. N. Bregman (2001) Chandra X-ray observations of the X-ray faint elliptical galaxy NGC 4697, *ApJ* **556**, 533.
- Schachter, J., A. V. Filippenko & S. M. Kahn (1989) Bowen fluorescence in Scorpius X-1, *ApJ* **340**, 1049.
- Schreier, E., et al. (1972) Evidence for the binary nature of centaurus X-3 from Uhuru X-ray observations, *ApJ* **172**, L79.
- Silva, D. R. & M. E. Cornell (1992) A new library of stellar optical spectra, *ApJS* **81**, 865.
- Stanimirovic, S., et al. (1999) The large-scale HI structure of the Small Magellanic Cloud, *MNRAS* **302**, 417.
- Stella, L., W. Priedhorsky & N. White (1987) The discovery of a 685 second orbital period from the X-ray source 4U1820-30 in the globular cluster NGC 6624, *ApJ* **312**, L17.
- Stephenson, C. B. & N. Sanduleak (1977) New H-alpha emission stars in the Milky Way, *ApJS* **33**, 459.
- Strohmayer, T. E. & C. B. Markwardt (2002) Evidence for a millisecond pulsar in 4U 1636-53 during a superburst, *ApJ* **577**, 337.
- Strohmayer, T. & L. Bildsten (2006) New views of thermonuclear bursts, in *Compact Stellar X-ray Sources*, 113, edited by W. Lewin & M. van der Klis, New York: Cambridge University Press.
- Strohmayer, T. E., et al. (1996) Millisecond X-ray variability from an accreting neutron star system, *ApJ* **469**, L9.
- Tananbaum, H., et al. (1972) Discovery of a periodic pulsating binary X-ray source in hercules from Uhuru, *ApJ* **174**, L143.
- Tauris, T. M. & E. P. J. van den Heuvel (2006) Formation and evolution of compact stellar X-ray sources, in *Compact Stellar X-ray Sources*, 623, edited by W. Lewin & M. van der Klis, New York: Cambridge University Press.
- Thorsett, S. E. & D. Chakrabarty (1999) Neutron star mass measurements. I. Radio pulsars, *ApJ* **512**, 288.
- van der Klis, M. (2005) Timing neutron stars, *AIPC* **797**, 345.
- van der Klis, M. (2006) Rapid X-ray variability, in *Compact Stellar X-ray Sources*, 39, edited by W. Lewin & M. van der Klis, New York: Cambridge University Press.
- van der Klis, M., et al. (1996) GX 5-1, *IAUC* **6511**.
- Verbunt, F. (1987) The formation of ultra-short period binaries in globular clusters, *ApJ* **312**, L23.
- Verbunt, F. (2005) X-ray sources in globular clusters, *AIP Conf Proc* **797**, 30.
- Verbunt, F. & P. Hut (1987) The globular cluster population of X-ray binaries, in *The Origin and Evolution of Neutron Stars*, 187, edited by D. Helfand & J.-H. Huang, Dordrecht: Reidel.
- Verbunt, F. & W. H. G. Lewin (2006) Globular cluster X-ray sources, in *Compact Stellar X-ray Sources*, edited by W. Lewin & M. van der Klis, New York: Cambridge University Press.
- Walter, F. M., et al. (1982) Discovery of a 50 minute binary period and a likely 22 magnitude optical counterpart for the X-ray burster 4U 1915-05, *ApJ* **253**, L67.
- Warner, B., P. A. Woudt & M. L. Pretorius (2003) Dwarf nova oscillations and quasi-periodic oscillations in cataclysmic variables - III. A new kind of dwarf nova oscillation, and further examples of the similarities to X-ray binaries, *MNRAS* **344**, 1193.
- White, N. E. & L. Angelini (2001) The discovery of a second luminous low-mass X-ray binary in the globular cluster M15, *ApJ* **561**, L101.
- Wijnands, R., et al. (2003) Quasi-periodic X-ray brightness fluctuations in an accreting millisecond pulsar, *Nature* **424**, 44.
- Woosley, S. E. & R. E. Taam (1976) Gamma-ray bursts from thermonuclear explosions on neutron stars, *Nature* **263**, 101.

Chapter 12

Black-hole X-ray binaries

12.1 Introduction

12.1.1 Concept of a black hole

Black holes have attracted people's imaginations perhaps more than any other kind of object in the cosmos. Remarkably the concept of a black hole dates back more than 200 years. In 1783, the Cambridge cleric John Michell speculated in a lecture to the Royal Society about the effects of the Sun's gravity on the light it was radiating. Michell was aware of the finite speed of light (determined by Roemer in the seventeenth century from observations of eclipse timings of Jupiter's moons) and believed that photons from the Sun (he called them 'corpuscles') would be slowed down as they left the Sun due to its gravity. His speculation was to point out that if the Sun's diameter were 500 times larger and of the same density, then its mass would be $10^8 M_\odot$, and gravity would prevent light from escaping the Sun at all. A similar conjecture was put forward by Laplace in 1795.

However, our modern concept of a black hole stems from Einstein's theory of general relativity (GR) and the first exact solutions, derived by Karl Schwarzschild in 1916, of Einstein's equations. Under GR, the effect of a massive body's gravity is to curve the space-time around it, forcing light to follow a curved path (called a 'geodesic'). If the body is sufficiently massive and compact, then this curvature closes in on itself, and any light emitted by the body will never escape – hence the term *black hole*.

The massive, dense body creating the black hole therefore has no visible surface, a

consequence which distinguishes black holes from other compact objects and may provide an observational signature by which they can be recognised. A well-defined region around a black hole is set by the radius at which the escape velocity is equal to the speed of light (i.e. $1/2mc^2 = GMm/R_S$) and hence $R_S = GM/c^2$. First defined by Schwarzschild, and written as R_S , it is better known by the emotive term the *event horizon* as no information about processes occurring within that radius can be transmitted across it.

It is therefore remarkable that we can obtain observational material on such extreme theoretical concepts at all. Indeed, for the first half of the twentieth century it was felt that they would remain more in the realm of mathematicians than physicists. But the discovery of pulsars and quasars brought such astrophysical exotica into mainstream physics, and with X-ray binaries we have the potential to directly determine their properties.

12.1.2 Black holes in binary systems

All the discussion of interacting binaries so far has been in terms of a 'compact object' which is accreting matter from its companion. In most cases this has been either a neutron star or a white dwarf. But the basic model of the binary is almost independent of what type of compact object it contains. As long as there is some idea of its distance, then the nature of the compact object can usually be inferred from the observed X-ray luminosity. Because the gravitational potential of a neutron star is so much deeper than that of a white dwarf (by a factor of 1000), it can produce

that much more luminosity from a given amount of accreted matter. However, the accretion rates can vary dramatically (e.g. when an X-ray binary enters a ‘quiescent’ state, in which mass transfer onto the compact object is limited for some reason), and discrimination between different types can then be quite difficult. Fortunately, neutron stars sometimes signal their presence in distinct and unmistakable ways, particularly through very precise X-ray pulsations (as a result of the neutron star’s spin) and type I X-ray bursts (due to thermonuclear burning on the neutron star’s surface), as described in Chapter 11.

How, then, can black holes in X-ray binaries be identified in a manner that yields an unambiguous result, especially since the difference between the gravitational potential for a neutron star and for a stellar-mass black hole is much smaller? One way is to directly measure the mass of the compact object.

It is known that there must be a limit to the mass of a neutron star and that this limit is almost certainly in the region of $3 M_{\odot}$. Most theorists are convinced that it is somewhat less than $3 M_{\odot}$, but this continues to be intensely debated, and observers are actively seeking evidence for ‘heavy’ neutron stars. Some theorists have argued that it is not necessary to know the equation of state of the material inside a neutron star (about which much has been inferred from observations of radio pulsars); rather one can simply assume that GR is correct and that causality is obeyed (i.e. the requirement that the speed of sound inside the star is less than the speed of light). From these two assumptions, theorists have shown that a (non-spinning) neutron star cannot be bigger than $3.2 M_{\odot}$. Allowing for rotation, this value is increased by about 20 per cent. The observed masses of neutron stars, as derived from those binaries containing X-ray and radio pulsars, are all consistent, with values substantially below $3 M_{\odot}$, with the average being remarkably close to the canonical value of $1.4 M_{\odot}$ (see Fig. 11.12).

12.1.3 Black holes represent the ultimate test for physics

Massively dense objects are very important to twenty-first-century physicists. The equation of state of matter at such densities and the very

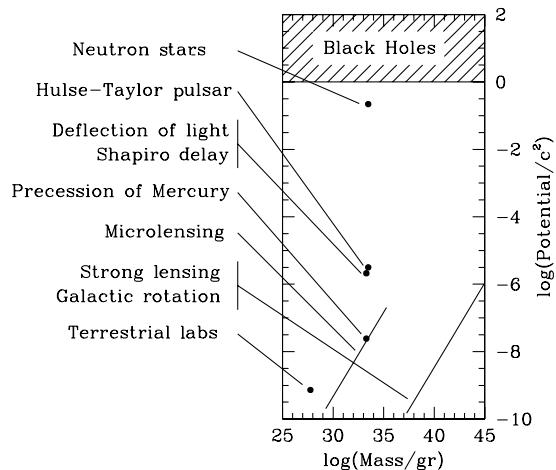


Fig 12.1 How close have we come to testing general relativity in the extreme limit? From Earth-based experiments to observations close to neutron stars and black holes, the potential energy release is plotted here as a fraction of the maximum achievable (Psaltis, 2006).

nature of what would be happening inside a black hole could not be inferred by observation in any other way. Indeed, our current physical laws are based on Einstein’s theories (of relativity), and yet these have only been tested in cases in which the perturbations to Newton’s laws are really quite small, as demonstrated in Figure 12.1. For any particle of mass m at distance R from a star of mass M , its potential energy is GMm/R , whereas the maximum energy that can be extracted from the particle is $E = mc^2$. Hence the fraction of this maximum that is emitted upon accretion is simply $G M / R c^2$, and it only approaches unity in the vicinity of a neutron star’s surface or the event horizon of a black hole. Note that this fraction depends only on the two basic properties, M and R , of the compact object. Even the binary radio pulsar discovered by Hulse and Taylor is only testing GR at a level a factor of 1 million below the maximum! Consequently, the environment of a black hole, with the most extreme gravitational fields, represents the ultimate test of GR.

The evidence for the very existence of black holes must then come from accurately determining the masses of as many compact objects as possible and then finding and studying those that are clearly above the maximum mass that could be

sustained as a neutron star. Such work is mostly confined to objects in our Galaxy but has recently been extended to include some of the most luminous objects in the Local Group of galaxies (in particular, the Magellanic Clouds, M31 and M33). The very presence of luminous X-rays from a galactic binary tells us that we must be dealing with a compact object. A high luminosity takes it further by limiting the choice solely to neutron stars and black holes. Can we therefore determine the mass of the compact object by direct measurement alone, and what are the limitations of this approach? And having done so, can we use the common observed properties of the black hole systems to be able to infer the presence of black holes in others?

12.2 | Measuring masses

The most direct measurement comes from optical spectroscopy, which gives the orbital radial velocity curve of either the secondary (mass losing) star or, if there is emission present, the accretion disc around the compact object. From these measurements alone, the mass function (as described in Box 11.4) can be obtained. Ideally, the compact object itself can be followed if, for example, it is an X-ray pulsar whose period variations define the orbital motion directly (see Figure 11.10). However, there is one fundamental difference with potential black-hole systems in that we cannot measure K_X (the velocity amplitude of the compact object) in this way because, by their very nature, the black holes cannot provide a periodic signature (such as a pulsation) to be followed around the orbit. Indeed, the presence of a pulsation tells us at once that we cannot be observing a black hole and that the compact object must be a neutron star. For a black hole, such a signature can only be produced inside the event horizon, which we cannot observe. Instead, we are measuring K_2 (the velocity amplitude of the mass donor, which is actually $V_2 \sin i$, where V_2 is the true orbital velocity and i is the inclination of the orbital plane to our line of sight), and so the mass function is

$$f(M) = \frac{P K_2^3}{2\pi G} = \frac{M_X^3 \sin^3 i}{(M_X + M_2)^2}, \quad (12.1)$$

which depends only on the directly observable quantities P and K_2 (frequently i is poorly known). The limitations are clear. If M_2 is large (as in the HMXBs Cyg X-1 and LMC X-3), then the mass function is a small number. To derive a reliable value for M_X , we must know M_2 and i accurately. M_2 can only be inferred from the optical spectral type of the star and our knowledge of the masses of other stars of the *same* spectral type. Unfortunately, such studies, which have produced accurate masses for normal stars, are confined to nearby, *non-interacting* binaries, whereas a black-hole candidate, again by definition, is a very close, *interacting* binary in which large amounts of mass have been transferred between the components over the binary's lifetime. They have therefore reached their current configuration by very unusual evolutionary processes (as described earlier in Chapter 11), and this introduces the potential for large systematic uncertainties in the mass estimation.

However, if M_2 is small (i.e. the system is an LMXB), then the mass function itself represents a firm *lower limit* to the value of M_X (if $M_2 = 0$ and $i = 90^\circ$ in the equation, then $f[M] = M_X$). Obviously, M_X must be *larger* than this when realistic values for M_2 and i are inserted.

12.3 | Observed properties

One of the brightest X-ray sources in the sky is Cyg X-1, known from the early days of X-ray astronomy and still reknowned as the 'first' black-hole system to be found in our Galaxy. Indeed, its association in 1971 with an eighth magnitude, hot, blue star (HDE 226868), allowed Webster and Murdin (1972) to demonstrate its binary nature, a seminal result in X-ray astronomy. This was the first clear demonstration that binary systems were involved in producing luminous X-ray sources. Furthermore, optical spectroscopy revealed the amplitude of motion of the bright star (almost 50 km s^{-1}) and hence allowed the first estimate of the X-ray-emitting object's mass to be made, as described earlier. When combined with its period of 5.6 days, this mass was calculated to be $> 15 M_\odot$ (using the technique described in Box 11.4). This is well outside any conceivable mass for a neutron star, hence the

Table 12.1 Observed properties of black-hole X-ray binaries.^a

X-ray source	Optical/IR counterpart	Year ^b	Spec.	P_{orb} (hours)	$f(M)$ (M_{\odot})	M (M_{\odot})
<i>LMXBs</i>						
GRS1915+105	V1487 Aql	1992/Q ^c	K/MIII	804.0	9.5 ± 3.0	10.0–18.0
J1859+226	V406 Vul	1999/1	–	9.2	7.4 ± 1.1	7.6–12.0
J1550–564	V381 Nor	1998/5	G8/K8IV	37.0	6.86 ± 0.71	8.4–10.8
J1118+480	KV UMa	2000/2	K5/M0V	4.1	6.1 ± 0.3	6.5–7.2
GS2023+338	V404 Cyg	1989/1 ^d	K0III	155.3	6.08 ± 0.06	10.1–13.4
GX 339–4	V821 Ara	1972/>10	–	42.1	5.8 ± 0.5	–
GS1354–64	BW Cir	1987/2	GIV	61.1	5.75 ± 0.30	–
GS2000+251	QZ Vul	1988/1	K3/K7V	8.3	5.01 ± 0.12	7.1–7.8
H1705–250	V2107 Oph	1977/1	K3/7V	12.5	4.86 ± 0.13	5.6–8.3
GRS1009–45	MM Vel	1993/1	K7/M0V	6.8	3.17 ± 0.12	4.6–5.8
J1124–684	GU Mus	1991/1	K3/K5V	10.4	3.01 ± 0.15	6.5–8.2
J1650–500		2001/1	K4V	7.7	2.73 ± 0.56	–
A0620–003	V606 Mon	1975/1 ^d	K4V	7.8	2.72 ± 0.06	8.7–12.9
J0422+32	V518 Per	1992/1	M2V	5.1	1.19 ± 0.02	3.7–5.0
<i>HMXBs</i>						
J1819.3–2525	V4641 Sgr	1999/4	B9III	67.6	3.13 ± 0.13	6.8–7.4
J1655–40	V1033 Sco	1994/3	F3/F5IV	62.9	2.73 ± 0.09	6.0–6.6
LMC X-3		–	B3V	40.9	2.3 ± 0.3	5.9–9.2
M33 X-7		–	O7III	82.9	0.46 ± 0.8	14.2–17.1
4U1543–475	IL Lup	1971/4	A2V	26.8	0.25 ± 0.01	8.4–10.4
Cyg X-1	HDE 226868	–	O9.7Iab	134.4	0.244 ± 0.005	6.8–13.3
LMC X-1		–	O7III	93.8 ^c	0.13 ± 0.05^c	4.0–10.0

^a from Remillard and McClintock (2006) and Charles and Coe (2006).

^b Year of initial X-ray outburst/total number of X-ray outbursts.

^c ‘Q’ denotes quasi-persistent intervals (e.g. decades) rather than typical outburst.

^d Additional outbursts in optical archives: A 0620 (1917) and V404 Cyg (1938, 1956).

implication that it must be a black hole, and Cyg X-1 rapidly became synonymous with ‘black hole’. A second object of this type, LMC X-3, was identified in the 1980s. Together with Cyg X-1, these are still considered to be among the best black-hole candidates in the HMXB class. Their properties are summarised in Table 12.1 and shown schematically to scale in Colour Plate 33.

However, the fundamental limitation in this analysis (and which is appealed to by all who argue against the existence of black holes in these systems) is that the mass of the compact object depends on knowing the mass of the mass donor, and as described in the previous section, this has to be inferred from optical spectroscopy. The mass

of the compact object is pinned to the mass of the mass-losing star, which is determined from its observed spectral type. The problem then is, why should a star of a given apparent spectral type in an X-ray binary be of the same mass as a similar star in a non-interacting binary? Since mass loss must be taking place (onto the putative black hole), then this assumption is indeed questionable.

Table 12.1 indicates that LMC X-3 might be one of the better candidates for a black hole. After all, it does have quite a large mass function, even though it has an early-type spectrum which implies the presence of a massive secondary star, akin to Cyg X-1. Also, being in the Large

Magellanic Cloud [LMC], one of our nearest neighbouring galaxies), its distance of 55 kpc is more accurately known than many galactic X-ray binaries. The observations infer a massive compact object of at least $6 M_{\odot}$, again well in excess of that considered plausible for a neutron star. However, this interpretation has also been challenged. The mass estimates are based on the assumption that all the optical light is coming from the secondary star. Whilst reasonable in the case of Cyg X-1 (which is an extremely bright supergiant star), it is clear from more detailed observations that there is evidence for an accretion disc contributing to the light seen from LMC X-3. Allowing for this, the mass of the secondary star could be as little as $0.7 M_{\odot}$ and that of the compact object only $2.5 M_{\odot}$. (In this situation, the reason the optical star appears like the much more massive B3V star is because we are seeing the hot core of an evolved star, sometimes called an OB subdwarf and not a large, early-type main sequence star.) Clearly interpreting the observations of compact objects with large, or apparently large, companion stars is not at all straightforward.

12.4 Soft X-ray transients (or X-ray novae)

Transient X-ray sources were introduced earlier (Chapter 11) and have been known from the early days of X-ray astronomy. Indeed, they represent approximately half the known X-ray binaries in our Galaxy, and they clearly divide into two classes. Those associated with the massive Be stars (section 11.2.4) have hard (hot) X-ray spectra, making them very different from the soft X-ray transients, which contain a low-mass donor. In most other respects the soft X-ray transients, also known as X-ray novae (XRN)¹ are indistinguishable from the low-mass X-ray binaries described in detail in Chapter 11, but for a variety of reasons the mass transfer onto the compact object is highly erratic. Some of these XRN (roughly a quarter) are neutron star LMXBs, but it is of great

interest for black-hole studies that the majority of XRN are proving to be very good *black-hole candidates* (or BHC).

For typically years at a time, most XRN are quiescent (yielding very weak X-ray fluxes only detectable by XMM or Chandra) and optically faint (typically 20th magnitude or fainter). But then the mass transfer onto the compact object turns on, and a powerful X-ray outburst occurs, many of which have now been picked up by the all-sky monitor (ASM) of NASA's Rossi X-ray Timing Explorer (RXTE), from which a sample set of XRN light curves considered here are plotted in Fig. 12.2. At this point the accretion disc brightens dramatically, sometimes by as much as 7 or 8 magnitudes. Indeed, at this brightness, they can be observable by amateur variable star observers, and one such object (V404 Cyg) was first noted 70 years ago. The peak brightness is reached in a matter of days, and then a gradual decline sets in that can take months.

The third source in our slate of historically significant BHC is A0620-00, which has become the archetypal soft X-ray transient. This source was, for a time, 3 times as bright as Sco X-1, normally the brightest source in the sky. A0620-00 was discovered by Elvis *et al.* (1975) when it flared on 3 August 1975 and was immediately detected by both Ariel V and SAS-3, two spacecraft which provided a long-term monitoring capability. The outburst lasted for 8 months, and its overall X-ray light curve is shown in Fig. 12.3. During the subsequent 30 years, an additional 50–60 X-ray transients have been observed, some with very similar light curves to A0620-00.

The optical counterpart of A0620-00, designated V616 Mon, was rapidly identified through its novalike behaviour (it had brightened by over 7 magnitudes from its quiescent 18th magnitude). Once it had returned to quiescence, it was found to consist of a late-type (K5V) star bound to its (optically) unseen companion in a 7.8-hour binary. Although A0620-00 does not eclipse, McClintock and Remillard (1986) measured a very high velocity amplitude which inferred a compact object mass of at least $3.2 M_{\odot}$ (for an inclination of 90°).

¹ Not to be confused with *classical novae*, which are caused by irregular thermonuclear eruptions on the surface of an accreting white dwarf in a low-mass binary system.

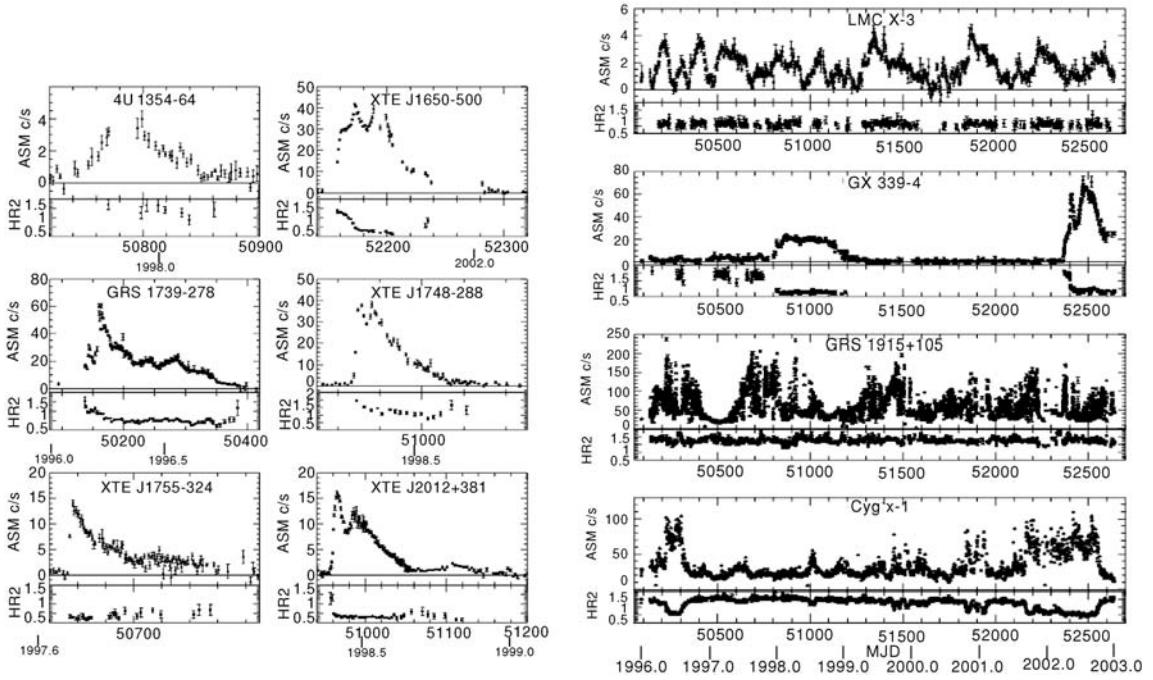


Fig I2.2 Long-term (7-year) X-ray light curves of (left) six transient black-hole binaries and (right) four steady black-hole binaries. The lower panels show the spectral variations via the X-ray hardness ratio (adapted from McClintock & Remillard, 1986).

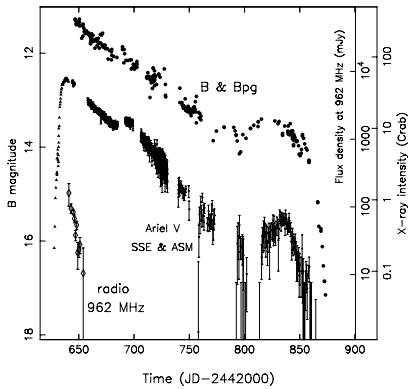


Fig I2.3 The outburst light curves at (top to bottom) optical, X-ray and radio wavelengths of one of the brightest soft X-ray transients ever observed, A0620-00, in 1975.

Reasonable assumptions for the inclination and mass of the normal star will lead to masses for the compact object greater than this. As explained in Section 12.2, this limit is much firmer than the mass constraints for the HMXBs.

12.4.1 The distorted secondary star

There is a further constraint imposed by the observation of geometrical distortions in the secondary star, which are caused by the gravitational field of the compact object. A Roche geometry star (see Box 11.2) presents a varying size to the observer during its orbital cycle, and so the star is seen to vary slightly in brightness (as shown schematically in Fig. 12.4). Modelling this effect (an example is shown in Fig. 12.5) provides an estimate of the binary inclination i , it being clear from the schematic that, viewed pole-on, there would be no modulation seen at all, whereas in the orbital plane the modulation is at a maximum.

Armed with a value for i allows a better estimate of the compact object mass to be made. Indeed, the lower limit on the mass is then raised to at least $7 M_{\odot}$, but this value is now more model-dependent than the firm limit given earlier. Nevertheless, A0620-00 is a much stronger BHC than Cyg X-1 and LMC X-3 because there are fewer questionable assumptions.

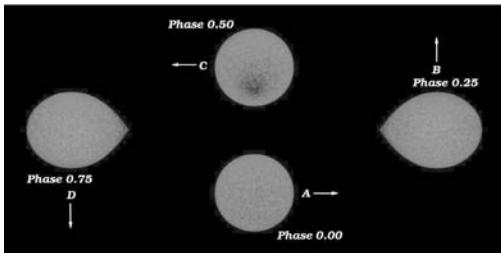


Fig 12.4 Illustration of the ellipsoidal variations in a close binary. The powerful gravitational influence of the compact object distorts the shape of its mass-losing companion. It adopts the shape of the Roche lobe (see Chapter 11). The net result of this is that the light output of the star varies with binary phase because of the variation in radiating area around the orbit. These effects are easiest to detect in the red, where the companion star's light usually dominates.

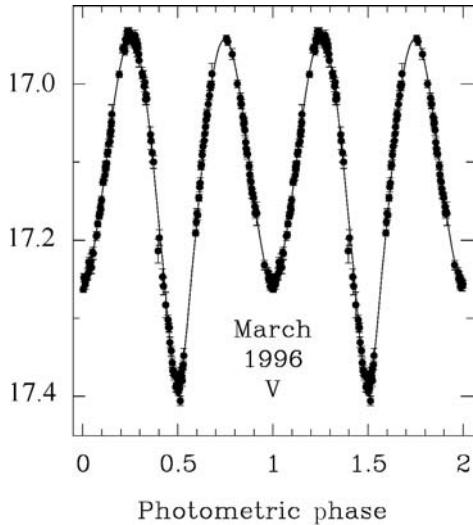


Fig 12.5 Observed orbital light curve in quiescence of the X-ray transient GRO J1655-40, a quite superb example of the ellipsoidal modulation of the mass donor shown earlier. Note how this leads to a modulation with two maxima and two minima for each orbital cycle (Orosz & Bailyn, 1997).

However, those working in this field sought a stellar-size object whose minimum mass, defined from the most straightforward and unambiguous observations, clearly excludes the possibility of its being a neutron star. This *holy grail* of an object had to have a mass function of at least $5 M_{\odot}$, putting it clearly outside the maximum mass possible for a neutron star. Whilst putting the finishing touches on the first edition of this book,

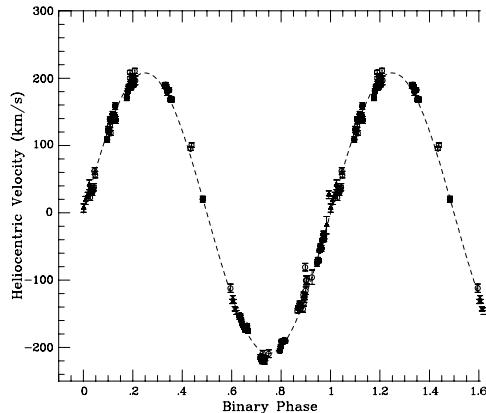


Fig 12.6 Radial velocity curve showing the motion of the companion star in the V404 Cyg binary system. These measurements were made from optical spectroscopy using the 4.2-m William Herschel Telescope on La Palma; the solid curve is that expected for a circular orbit. The amplitude of this motion is very high (over 400 km s^{-1} peak to peak) and gives a very strong example for the existence of galactic black holes (Casares & Charles, 1994).

one of us (P.A.C.) together with colleagues Jorge Casares and Tim Naylor found such an object. Again, it was an XRN, discovered during an outburst in May 1989 by the Ginga X-ray satellite and designated GS2023+338. This was quickly identified with a previously known recurrent nova, V404 Cyg, which was first recorded in 1938. It has flared several times since then (optically), but its unusual nature was not recognised until the enormous X-ray output was detected by Ginga.

Such an X-ray flux (almost as bright as Sco X-1) showed that it could not be a typical recurrent nova (which is actually a type of cataclysmic variable; see Chapter 10) but must at least involve a neutron star. And its change in brightness of more than 7 magnitudes from quiescence to outburst showed that it must be an LMXB, similar to A0620-00. Consequently, when it returned to quiescence in the early 1990s, spectra of V404 Cyg revealed the late-type (K0) secondary star. More importantly, the absorption lines displayed large radial velocity variations (amplitude 210 km s^{-1} ; see Fig. 12.6) on a period of 6.5 days, at that time the longest period in the class (and still the second longest). Using these two directly observed parameters gives a mass function of $6.3 M_{\odot}$. This was the

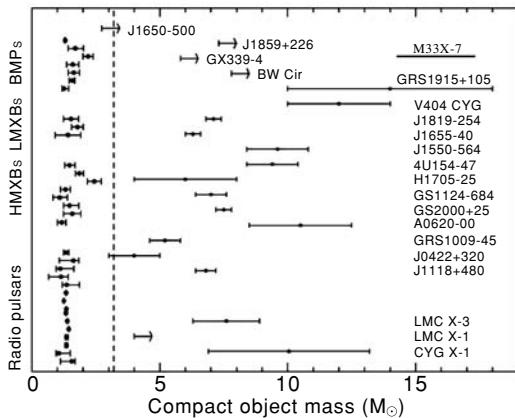


Fig 12.7 Dynamically measured masses for known neutron stars and black-hole candidates. It is quite remarkable how the neutron star masses (left of dashed line) cluster so close to the canonical value of $1.4 M_{\odot}$, whereas the BHC cover a much wider range.

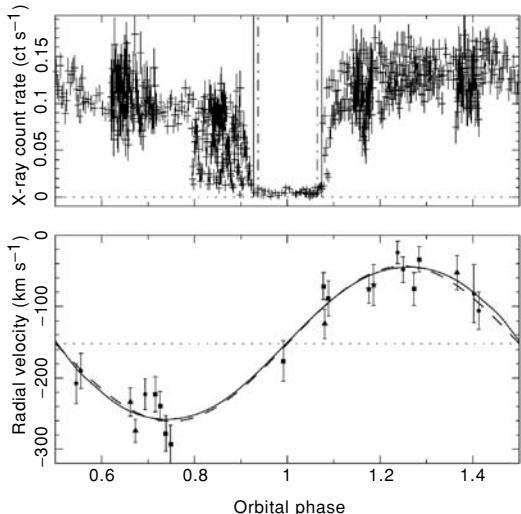


Fig 12.8 (top) X-ray light curve (Chandra ACIS) of M33 X-7 as a function of orbital phase. The eclipse is marked with two vertical solid lines which include the effect of the stellar wind. The dash-dotted lines show the extent of the photosphere. (bottom) Orbital radial velocity curve (data points) of the O star together with a model fit (solid line) and sinusoid (dashed line) (from Orosz et al., 2007).

absolute *minimum* mass for the compact object, and hence V404 Cyg represented a milestone in this field as a clearly unambiguous BHC.

Similar studies over the last decade of XRN discovered by RXTE have built up the number of BHC to those summarised in Table 12.1, the masses of which are plotted in Fig. 12.7.

12.5 Eclipsing black-hole binaries

As with the neutron star HMXBs described in chapter 11, the best systems for accurate mass determinations would be those in which the X-ray source is eclipsed, and hence the orbital inclination is tightly constrained. Unfortunately, none of the XRN described here are eclipsing. This is an observational selection effect that is also mentioned in chapter 11 because eclipsing LMXBs are much rarer due to the obscuring effects of the accretion disc itself; i.e. the mass donor in LMXBs subtends a much smaller angle at the X-ray source than is the case for short-period HMXBs, and so the chances of eclipse are correspondingly smaller. Much more likely would be the discovery of an eclipsing HMXB involving a black hole, but none are known in our Galaxy as the HMXB population is dominated by neutron star systems.

There was therefore considerable excitement when the first eclipsing black-hole binary was identified by Jerry Orosz and colleagues in the Local Group galaxy M33. It is a luminous X-ray source in M33, designated X-7, first discovered with the Einstein telescope, which was able to resolve it from nearby sources. An intrinsically bright optical counterpart has been identified, which means that it must be an HMXB as an LMXB would be unobservably faint at the distance (840 kpc) of M33. The observation of X-ray eclipses (Fig. 12.8) means that we (the observers) are close to the orbital plane and the orbital inclination must be close to 90° . So one of the big uncertainties in black-hole mass determination is now reduced.

The largest ground-based telescopes were employed to determine the radial velocity of the companion star to X-7 as a function of phase (Fig. 12.8). With an orbital period of 3.45 days, the duration of the eclipse and the known distance of M33 yield the size of the star, and the distance and spectrum give the luminosity. Thus the optical star is identified as an O7III star of mass $70 \pm 7 M_{\odot}$ (making it one of the most massive

stars whose mass is accurately known), and the compact object mass is deduced to be $15.4 \pm 1.4 M_{\odot}$, the highest yet which has been dynamically determined. A compact object with this mass must be a black hole, and furthermore, its precision is remarkable. Chapter 13 includes an XMM image of M33 which shows X-7.

12.5.1 Producing black holes in SN explosions

The significance of M33 X-7, and the few other stellar-mass BHs whose masses exceed $10 M_{\odot}$, becomes apparent when we think about how these objects must have been created. They can only have been the result of the stellar collapse of a very massive star, as described in Chapter 8. But calculations of the late stages in the evolution of such massive stars indicates that the pre-SN core becomes so hot that it literally blows off the outer layers of the star. We see these as *Wolf-Rayet* stars, whose optical and UV spectra are dominated by powerful emission lines indicating a fast, outblowing wind containing huge amounts of material (typical mass loss is $\sim 10^{-3}\text{--}10^{-4} M_{\odot} \text{ yr}^{-1}$, which even for a massive star cannot be sustained for very long). Consequently, by the time the star becomes a type Ib or II SN, it will have lost much of its original mass. Calculations of the residual mass available for the compact object, compared with the observed masses of galactic BHs, are displayed in Fig. 12.9.

Even with the relatively small number of dynamically determined compact object mass measurements, there are interesting differences between the calculated and observed distributions:

- As is also clear from Fig. 12.7, there is a dearth of measured masses in the $2\text{--}5 M_{\odot}$ range which is not reproduced in the evolutionary calculations; i.e. are there ‘heavy’ neutron stars or ‘light’ black holes? This is still an area of great debate, but there may be an observational selection effect here. Dynamical masses in LMXBs can only be obtained when the source is quiescent, i.e. the X-rays are off, thereby allowing the mass donor to be seen and studied. That is why this work has been concentrated in the X-ray transient category, but we may therefore

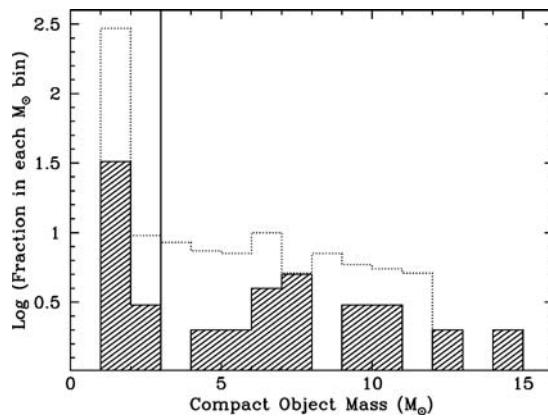


Fig 12.9 Dynamically determined black-hole masses (shaded histogram; from Table 12.1) compared with model calculations of residual compact object masses (dotted line) following WR-type mass loss and subsequent SN explosion (from Casares, 2007).

be selecting only objects in a particular mass range, although the reason for this is currently unclear.

- There is a cut-off expected above a certain mass, which the calculations suggest is around $12 M_{\odot}$. However, this is already being challenged by the observations, particularly the mass of GRS 1915+105, which with a low-mass companion could not have grown substantially since its formation. Note that while M33 X-7’s mass is even higher, it does have a very high mass companion, so large-scale mass exchange may have taken place.

12.6 Spectra of XRN

Luminous XRN containing black holes and neutron stars have characteristic X-ray spectra as shown in Fig. 12.10. Both require the presence of a thermal accretion disc component, which dominates at low X-ray energies (hence the ‘soft’ label), but the second components in their spectra are different:

- As demonstrated by Nova Mus 1991, the BH XRN usually display a power-law component, which can extend to extremely high energies (to at least the detector’s limiting energy).

Table 12.2 States of black-hole X-ray binaries.

State	Other name	Main component	Variability
Steep power law (SPL)	VHS	Power law	0.1–30 Hz QPOs
Thermal	HS	Thermal disc	No QPOs
Hard	LH	Power law	Possible QPOs

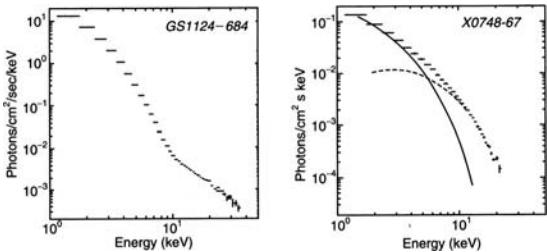


Fig 12.10 High-state X-ray spectra of (left) the BHC Nova Mus 1991 and (right) the X-ray burster X0748-676. The BHC spectrum is dominated at low energies by a thermal component but has an underlying hard power-law component that extends to very high energies. The X-ray burster, however, is well described by two thermal components, one from the disc, the other from the neutron star (adapted from Tanaka, 2008).

- The neutron star systems, however, require another thermal component, but this one is harder (i.e. hotter) than that of the disc. That this additional thermal component is well fit by a single temperature blackbody gives a strong clue as to its likely origin, namely, the surface of the neutron star itself. It is here that the accreted material comes to rest and deposits its kinetic energy.

12.6.1 X-ray spectral state transitions

Even in the early days of X-ray astronomy, Cyg X-1 had been noted as undergoing spectral transitions in X-rays which were associated with its radio emission. These X-ray spectral changes are easily visible in the 7-year light curve plotted in Fig. 12.2 (bottom right), where it switches between the HS (high, soft) and LH (low, hard) states.

These spectral states have usually been labelled according to whether they are ‘high’ or ‘low’ in brightness and ‘hard’ or ‘soft’ in spectral shape, as observed by the majority of X-ray astronomy

detectors over the last 30 years. However, with most X-ray astronomy satellites, this meant that the 2–20 keV range was dominant in deciding these labels. Once it became possible to regularly observe to much higher energies (e.g. with the Compton Gamma-Ray Observatory, and now INTEGRAL), it was clear that these labels were not necessarily accurate descriptions, e.g. GX339-4 in its LH (low, hard) state appears to CGRO as a brighter source than it does in its HS (high, soft) state! A clearer, simpler classification has been proposed recently by Remillard and McClintock (2006), as summarised in Table 12.2 (the old terms are also given due to the large volume of literature available that uses them), where the brightness criterion has been abandoned. Observed spectra (Fig. 12.11) of the XRN GRO J1655-40 in different states during its outburst are used to illustrate them.

A physical ‘visualisation’ of what might be happening in XRN to explain their transitions between these states was proposed by Ann Esin, Jeff McClintock and Ramesh Narayan (Fig. 12.12). It is based on the main assumption that, in the higher thermal states, the disc reaches closer to the compact object, whereas in the hard states, it is replaced by a much hotter but lower density medium. However, the range of behaviour is complex, and the literature on XRN X-ray spectral behaviour is large. For a much more detailed discussion, see Remillard and McClintock (2006).

12.6.2 Formation of jets in XRN: Microquasars

The similarities between some of the properties of AGN and X-ray binaries were well established by the discovery of the precessing jet source SS433 in the late 1970s (Chapter 11). But this analogy was taken to a new level with two XRN in the

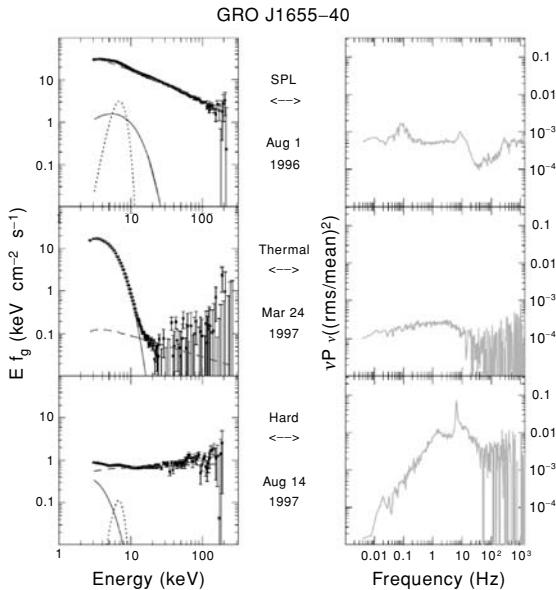


Fig 12.11 (left) X-ray spectra of the BHXB GRO J1655-40 obtained by RXTE, showing the three main activity states of XRN, from top to bottom, the very-high state (or ‘steep power law’, SPL), the high state (or ‘thermal’), and the low, hard state (or ‘hard’). The dominant spectral components in each state have been separated into the thermal (solid curve) and power-law (dashed curve) plots. Both are always present but in differing ratios. (right) Corresponding variability property of each state, plotted here as the power density spectrum. QPOs are evident in the SPL and hard states (from Remillard & McClintock, 2006).

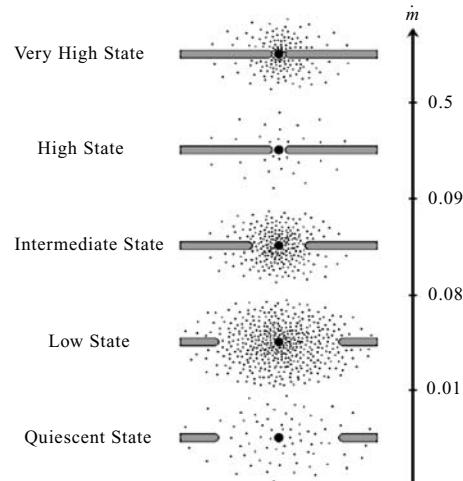


Fig 12.12 Sketch of how the accretion disc inner structure and flow varies amongst the different spectral states. The dotted regions represent ADAFs (see text) where the energy of the gas is quickly lost into the central black hole. The accretion rate (right axis) is estimated to vary over a range of more than a factor 100 from the brightest state (VHS) to quiescence (from Esin et al., 1997).

mid-1990s, GRO J1655-40 and GRS 1915+105, both of which displayed not only expanding radio jets (Fig. 12.13) but superluminal motion, as well, in a series of seminal radio observations by Felix Mirabel and Irapuan Rodrigues. Of course, such motions are not really superluminal and are simply an effect that occurs when something is moving rapidly towards or away from us, but at a small angle to our line of sight. This geometrical explanation (see Box 12.1) then allows us to infer the intrinsic speeds of the jet or blobs.

In the case of GRS 1915+105, there are clearly blobs moving in both directions, one towards us, the other away. The radio images of Fig. 12.13 give fractional velocities β for the two components as $\beta^+ = 0.64$ (receding) and $\beta^- = 1.49$ (approaching). From the equations in Box 12.1 we can calculate that both are consistent with a single β of 0.98 and

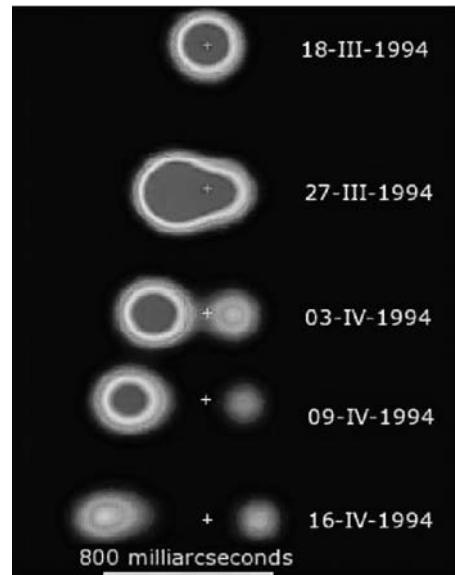


Fig 12.13 Radio ‘blobs’ ejected at apparent superluminal speeds from XRN GRS 1915+105 over a period of a month in 1994 (from Mirabel, 2007).

Box 12.1 | Superluminal expansion

An effect first discovered in radio QSOs, where emitting radio blobs were observed to be moving away from the central QSO at superluminal speeds, the explanation is a straightforward consequence of the finite speed of light, c . Consider a source which flares at a time t_1 and ejects a blob of material that moves away at velocity V , as shown in Fig. 12.14. After a time t_2 , the blob is observed again, when it is a distance D from the observer, and appears to him or her to have moved a distance on the sky of $V \Delta t \sin \theta$.

What will be the measured times of receipt of the photons emitted at events P_1 and P_2 ? It is necessary to add the light-travel times to t_1 and t_2 :

$$t_2^{\text{obs}} = t_2 + D/c,$$

and similarly,

$$t_1^{\text{obs}} = t_1 + (D - V \Delta t \cos \theta)/c.$$

$$\text{Hence } \Delta t^{\text{obs}} = \Delta t(1 - V \cos \theta)/c,$$

from which the apparent speed of the blob across the sky to the observer is simply

$$V^{\text{obs}} = V \sin \theta / (1 - V \cos \theta),$$

or expressing it as a fraction β of the speed of light,

$$\beta^{\text{obs}} = \beta \sin \theta / (1 - \beta \cos \theta).$$

If the blob is travelling at 95 per cent of the speed of light ($\beta = 0.95$) and at an angle θ of 15° to the observer, then $\beta^{\text{obs}} = 2.98$! Even higher apparent speeds can be attained if the jet or blobs are aimed closer to our line of sight.

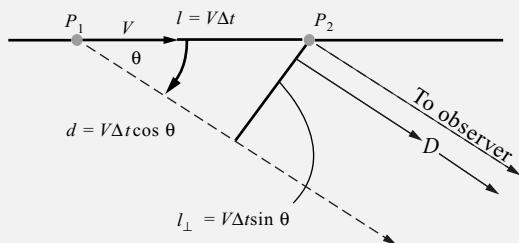


Fig 12.14 Schematic to illustrate how apparent superluminal motions can be observed. An event occurs in a source (P_1) at time t_1 that ejects a blob with velocity V in the direction shown, at an angle θ to the observer. A time t_2 later, the blob has moved a distance $l = V \Delta t$ away from the original source and is observed again (event P_2). To the observer, the apparent velocity is the projected distance moved on the sky, $V \Delta t \sin \theta$, divided by the time difference between receipt of the photons associated with events P_1 and P_2 .

an angle θ of 66° . The approaching blobs are the brighter ones (on the left in Fig. 12.13), a relativistic effect due to Doppler boosting in which the cone of emission is concentrated in the direction of motion. These effects are all identical to those observed in distant QSOs, hence the labelling of these XRN as ‘microquasars’.

The multi-wavelength light curves in Fig. 12.15 provide insight to the physical processes that must be taking place in the launching of these jets. The rapid X-ray variations are quenched within just a few minutes, following which there is an increase in the IR and then the radio. This disappearance of the inner accretion disc is followed by its replenishment and the associated launch of material in the form of a jet. As the jet material cools adiabatically, it radiates first in the IR, then in the radio, a process already well established in AGN

jets but on much longer timescales (years). To generate synchrotron radiation in the IR and up to X-ray wavelengths requires extremely energetic (TeV level) electrons to be accelerated close to the compact object during this process. Such energies have also been observed directly by ultra-high-energy γ -ray telescopes (e.g. HESS).

12.6.3 Unified model for state transitions and jet formation

With RXTE providing an abundance of high-quality X-ray spectra, rapid timing information and long-term (years) monitoring of these luminous X-ray binaries and XRN, attempts are being made to bring together the variety of X-ray states into a *unified model* that accounts for their observed behaviour. Many sources make spectral state transitions, and there are now ~ 20 microquasars

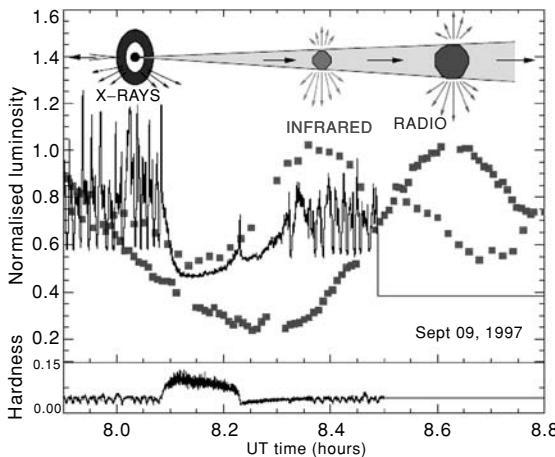


Fig 12.15 Subsequent X-ray, IR and radio light curves of GRS 1915+105 which show how the hot (X-ray) inner disc disappears and is immediately followed by the IR and radio emission from ejected material in the jet. The inner disc is then replenished, and the X-rays recover (from Mirabel, 1998).

where the radio properties appear linked to these spectral changes. Tommaso Belloni and Rob Fender have led a large team in producing the ‘unified model’ that is shown schematically in Fig. 12.16. The main panel in this figure plots the source’s X-ray intensity against its hardness ratio (simply a ratio of higher over lower energy channels so that a higher ratio means it is ‘hard’ and a lower ratio means it is ‘soft’) and is a convenient way of following how a source’s spectral state varies with time.

In a low-luminosity (<1% of maximum) hard (LH) state, steady radio jets are observed ([i] in Fig. 12.16), but as the mass accretion rate (and hence luminosity) increases (e.g. in an XRN outburst) and it makes a transition across to the thermal (HS) state, then transient jets can be seen ([ii] and [iii]). This is the type of behaviour that was displayed by GRS 1915+105 in Fig. 12.13. However, the range of details displayed by different sources is extremely complex, and this is a very active area of current research.

12.7 Black-hole spin

A stellar-mass black hole can be completely described by only two parameters; the mass M and angular momentum J . These two completely

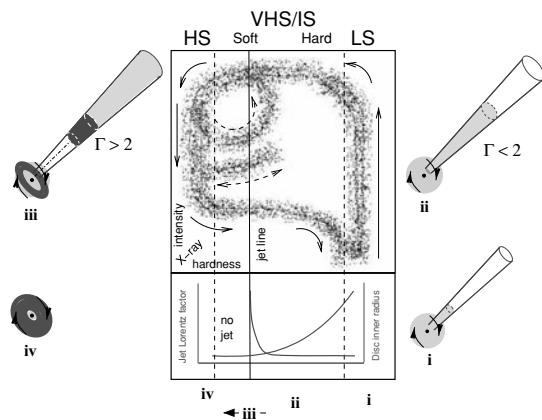


Fig 12.16 Schematic of how state transitions in XRN are related to the formation and quenching of jets. (top) X-ray intensity (y axis) against spectral hardness (x axis), with hardness increasing to the right. The tracks in this panel show how an XRN can move from one spectral state to another (LS = low/hard; VHS = very high; HS = high soft states), and it is observed that when crossing certain regions of this diagram, jets are formed or strengthened. (bottom) The jet’s Lorentz factor (speed) also plotted against hardness. The jet status at different points in these tracks is shown in the surrounding cartoons (from Fender et al., 2004).

determine the scale and nature of space-time surrounding the object. The maximum value possible for J is $J_{\max} = GM^2/c$. It is customary to refer to the spin of a particular object using a dimensionless spin parameter, a_* , which is simply the fraction of this maximum, i.e. $a_* = J/J_{\max} = Jc/GM^2$, and can range in value from -1 to 1 (negative values imply that it is counter-rotating with respect to the binary rotation); $a_* = 0$ corresponds to a non-spinning BH, which is referred to as a Schwarzschild BH, whereas $a_* = 1$ corresponds to a maximally spinning BH and is known as a Kerr BH (after the New Zealand mathematician who solved Einstein’s equations for this situation).

Orbiting accreting material moves through an accretion disc towards the star. However, the disc does not extend to the event horizon. Material cannot orbit closer than a radius called the innermost stable circular orbit (ISCO) as general relativity requires material closer than this to spiral rapidly towards the event horizon. A Schwarzschild (non-spinning) BH has $R_{\text{ISCO}} = 3R_S$, where R_S is the radius of the event horizon (the Schwarzschild Radius can be obtained by setting the nonrelativistic escape velocity = c , useful for memory but

not really appropriate for space-time close to the event horizon, i.e. $(0.5)mc^2 = GMm/R_S$, and hence $R_S = 2GM/c^2$). However, for a Kerr BH, $R_{ISCO} = 0.5R_S$, which means that the accretion disc can extend much closer to the event horizon (see Colour Plate 34).

J can be estimated by measuring the temperature of accreting material, which depends on R_{ISCO} . The closer the inner edge of the accretion disc is to the event horizon, the hotter the material. In practice, the X-ray spectrum is fit using model spectra. The model, which is fully relativistic, assumes a thin disc and blackbody emission with temperature dependent on radius. Statistically good data are available for several BH binaries, from which the spectral continuum has been well fit, allowing a_* to be calculated. Values obtained range from <0.25 for LMC X-3 to ≈ 1 for GRS 1915+105 (Narayan *et al.*, 2007), which has the highest disc temperature. A similar result is obtained for GRO J1655-40, which, together with GRS 1915+105, is a strong microquasar.

12.7.1 High-frequency QPOs and BH mass

The final property that links certain XRN to maximally spinning (i.e. Kerr) BHs is the existence of high-frequency QPOs. Their presence indicates that the disc must penetrate sufficiently close to the BH for stable orbits at those frequencies. And since the orbital frequency at R_{ISCO} for a Schwarzschild BH is $2200/M_{BH}\text{Hz}$ (and even higher for a Kerr hole), then such QPOs are providing a detailed probe of an extreme region close to the BH.

The fundamental oscillations possible in such a disc are complex. Resonance modes are expected to provide double-peaked QPOs with frequencies in the ratio of 3:2, and such pairs of QPOs have been definitely seen in at least three XRN, with a relation between the frequency and measured BH mass (Fig. 12.17). This opens the possibility of using this result in future observations to infer the compact object mass on the basis of its X-ray variability properties.

12.8 | How to pick out black holes?

A natural approach to take in identifying other BHC was to look for X-ray sources showing some

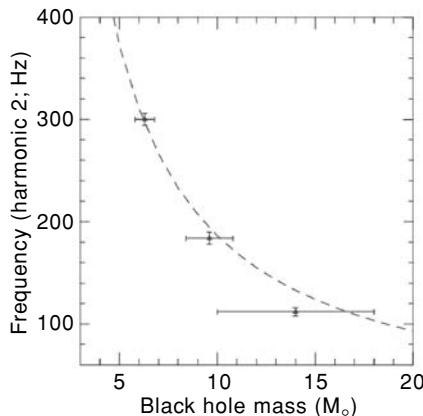


Fig 12.17 Plot of the link between high-frequency QPOs and BH mass for those XRN with dynamically determined masses (from left to right, XTE J1550-564, GRO J1655-40, GRS 1915+105). The observed QPOs are doubled, and the stronger QPO frequency is plotted here, which is twice the fundamental (from Remillard & McClintock, 2006).

of the characteristic properties of Cyg X-1 such as short timescale (millisecond) flickering (or shot noise). Subsequently, the (at times) soft X-ray spectrum of Cyg X-1 and many XRN was considered another potential black-hole property. Unfortunately, other sources which showed the flickering and soft spectra have now been ruled out as black holes on the basis of regular pulsations or X-ray bursts being subsequently discovered from them (see Chapter 11). The pulsations can only arise in a magnetosphere, which is not present in a black hole, and X-ray bursts require a neutron star surface. Flickering and soft X-ray spectra are therefore *not* definitive on their own but instead suggest that the object may be worthy of further study.

12.8.1 Direct evidence for the existence of an event horizon?

Much better would be to obtain a clear-cut method for confirming the presence of an event horizon. It has already been noted that the thermal state spectra of black-hole and neutron-star XRN are different in that the latter show two thermal components, with a hotter one that is interpreted as arising on the NS surface, there being no equivalent for the BH as the event horizon is not a physical surface. However, the spectra are very

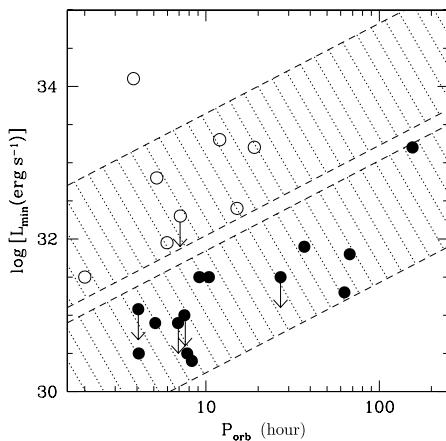


Fig 12.18 Comparison of the quiescent X-ray luminosities of NS (open circles) and BH (filled circles) XRN. The latter are much fainter, implying the existence of an event horizon (from Narayan & McClintock, 2007).

model-dependent, and this result is considered circumstantial. A more direct demonstration has been proposed by Ramesh Narayan, Mike Garcia and Jeff McClintock by comparing the quiescent X-ray luminosities of BH and NS XRN (Fig. 12.18). In this figure it is immediately apparent that the quiescent NS systems (open circles) are much more luminous (a factor of 10–100) than the BH systems with the same orbital period.

In the quiescent state, the accretion rate is low and the inner disc is expected to become much more tenuous, adopting what is referred to as an *advection-dominated accretion flow* (or ADAF). Such a flow has a hot, low-density gas which cannot efficiently radiate its energy before falling onto the compact object. This does not matter in the NS case as this energy will simply be transferred to the NS surface, which will be heated up and subsequently radiate. However, in the BH case, the hot gas has crossed the event horizon, and its energy is lost to the observer. Assuming that the mass transfer rates are similar in the two types for the same orbital period, then the main difference arises from the lack of a surface in the BH systems and thereby accounts for the observed luminosity differences. It is a powerful argument in favour of the existence of an event horizon in these XRN.

12.9 | Concluding remarks

The study of the BHC has shown how important it is to marry X-ray astronomy with other branches of astronomy. The prodigious X-ray output, combined with certain peculiarities of behaviour, gives us clues, but the accurate masses have all come from optical spectroscopy of the mass-losing companions. But it must be remembered that these masses do not tell us that a black hole is present. The argument is indirect. If the compact object has a mass of, say, 8 solar masses, then we infer the presence of a black hole because (1) we do not believe that neutron stars this massive can exist and (2) relativity tells us that a black hole is the only alternative. It is for this reason that scientists strictly refer to them as BHC. The search for black holes will therefore always be a difficult one, but it offers the exciting prospect of observationally approaching the frontiers of physics. Indeed, observations are becoming more precise and more dependent on the predictions of general relativity.

References

- Casares, J. (2007) Observational evidence for stellar-mass black holes, *IAU Symp* **238**, 3.
- Casares, J. & P. A. Charles (1994) Optical studies of V404 Cyg, the X-ray transient GS 2023+338, *MNRAS* **271**, L5.
- Charles, P. & M. J. Coe (2006) Optical, UV and IR observations of X-ray binaries, in *Compact Stellar X-ray Sources, Cambridge Astrophysics Series*, vol. **39**, 215 edited by W. Lewin & M. van der Klis, Cambridge: Cambridge University Press.
- Elvis, M., et al. (1975) Discovery of powerful transient X-ray source A0620-00 with Ariel V Sky Survey Experiment', *Nature* **257**, 656.
- Esin, A. A., J. E. McClintock & R. Narayan (1997) Advection-dominated accretion and the spectral states of BHXRBS, *Appl J* **489**, 865.
- Fender, R. P., T. M. Belloni & E. Gallo (2004) Towards a unified model for black hole X-ray binary jets, *MNRAS* **355**, 1105.
- McClintock, J. E. & R. A. Remillard (1986) The black hole binary A0620-00, *Appl J* **308**, 110.
- Mirabel, I. F. (2007) Black holes: From stars to galaxies, *IAU Symp* **238**, 309.

- Narayan, R. & J. E. McClintock (2008) Advection-dominated accretion and the black hole event horizon, *New Astron Rev* **51**, 733.
- Narayan, R., J. McClintock & R. Shafee (2007) Estimating the spins of stellar-mass black holes by fitting their continuum spectra, *AIPC* **968**, 265.
- Orosz, J. A. & C. D. Bailyn (1997) Optical observations of GRO J1655-40 in quiescence. I. A precise mass for the black hole primary, *Appl J* **477**, 876.
- Orosz, J. A., et al. (2007) A 15.65-solar-mass black hole in an eclipsing binary in the nearby spiral galaxy M 33, *Nature* **449**, 872.
- Psaltis, D. (2006) Accreting neutron stars and black holes, in *Compact Stellar X-ray Sources, Cambridge Astrophysics Series*, vol. 39, 1 edited by W. Lewin & M. van der Klis, Cambridge: Cambridge University Press.
- Remillard, R. A. & J. E. McClintock (2006) X-ray properties of black-hole binaries, *Ann Rev Astron Astrophys* **44**, 49.
- Tanaka, Y. (2008) Black-hole binaries, in *The Universe in X-rays*, 237, edited by J. E. Trümper & G. Hasinger, Berlin: Springer.
- Webster, B. L. & P. Murdin (1972) Cygnus X-1 – a spectroscopic binary with a heavy companion? *Nature* **235**, 37.

Chapter 13

Normal and starburst galaxies

13.1 X-ray sources in the Milky Way

Our Galaxy, the Milky Way, is a ‘normal’ spiral galaxy. It does not have a currently active nucleus (see Chapter 14), nor is it unusually luminous at any wavelength. Since we live in it, we find this a pleasing situation.

The first X-ray sources discovered within our Galaxy were naturally the brightest: the accretion-powered binaries. As time progressed, other Milky Way objects were also found to emit X-rays. Most of the emission from our Galaxy seemed to be from discrete objects, not from some galactic-sized diffuse region. This is also true for most other ‘normal’ galaxies. The populations of discrete sources in other galaxies is a major topic of interest. For any particular class of source we would like to know how the number of sources varies with X-ray luminosity. This function, the XLF, can be compared with theoretical results concerning the nature and evolution of the sources.

The most luminous galactic X-ray sources are the accretion-powered binaries, which consist of a compact object and a companion star. If the companion is spectral type A or later (mass $<\sim 1 M_\odot$), the system is a low-mass X-ray binary or LMXB. If the companion is spectral-type O or B (mass $>\sim 10 M_\odot$), the system is a high-mass X-ray binary or HMXB (Chapter 11).

Luminosities range from very low up to $\sim 10^{38}$ ergs s $^{-1}$. Those with high luminosity are thought to be operating close to the Eddington

limit, where the pressure of in-falling material is balanced by the pressure of outflowing radiation. Some binaries are extremely variable, even transient and undetectable when quiescent. Many LMXBs are within globular clusters where the high stellar density is favourable for the formation of exotic, compact binary systems (Chapter 11).

Cataclysmic variables (CVs) and supernova remnants (SNRs) are also strong X-ray sources with X-ray luminosities in the range $10^{35}\text{--}10^{37}$ ergs s $^{-1}$. CVs are accreting binary systems consisting of low-mass normal stars and white dwarf companions (Chapter 10). Many SNRs also contain isolated neutron stars and pulsar-wind nebulae (PWN). X-ray luminosities range from 3×10^{37} ergs s $^{-1}$ from the PWN within the Crab Nebula down to 10^{32} ergs s $^{-1}$ from PSR 1929+10, a nearby radio pulsar (Chapter 9). The largest class of galactic sources are stars, some perhaps binary but without compact companions. Observed coronal X-ray luminosities of stars range from 10^{33} to 10^{26} erg s $^{-1}$, the emission from the Sun (Chapters 6 and 7).

A large uncertainty in the study of galactic sources is the distance, which can often be uncertain by a factor of 2 or more, resulting in sometimes an order of magnitude uncertainty in the X-ray luminosity. Thus the luminosity function of the various types of galactic sources is unknown, and some interesting questions remain unanswered, e.g. are all bright accretion-powered binaries $1.4 M_\odot$ neutron stars operating at the Eddington limit, or are some of these black holes with larger masses?

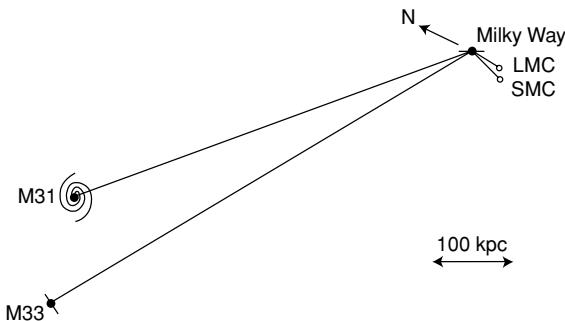


Fig 13.1 A scale drawing showing shape and separation of the principal members of the Local Group. These are projected on a plane containing the Earth's axis of rotation and oriented perpendicular to the plane of the galaxy (the Milky Way). All galaxies shown are approximately in this plane.

A nearby galaxy provides a large sample of X-ray sources, all at the same distance, so this uncertainty is eliminated. Also, since observations are made through a lesser amount of interstellar material, soft X-rays are detected that are unobservable from distant sources in the galactic plane. To observe, one needs a detector sensitive enough to find weak sources and with sufficient resolution to isolate them from their neighbours in crowded fields. Imaging X-ray telescopes can do this.

13.2 The Milky Way and the Local Group

The Local Group consists of two large stellar aggregations plus about 40 dwarf galaxies (Fig. 13.1). Most of the mass is in the Milky Way and in the nearby spiral galaxy, M31. M31 was long thought to be larger than our Galaxy, with 1.5 times as many stars, but recent data indicate that the two are about the same size. Another small galaxy, M33, is a face-on spiral and a bit more distant than M31. Our nearest neighbour, the Large Magellanic Cloud (LMC), is 50 kpc distant and contains about one tenth the number of stars in the Milky Way. Slightly farther away, the Small Magellanic Cloud (SMC) is about half this size. M31 and M33 are about 10 times farther away. The next nearest neighbours lie in groups ≈ 2.8 Mpc distant.

The Magellanic Clouds are visible to the naked eye but can only be seen clearly from the South-

ern Hemisphere. Both, with considerably less mass and irregular morphologies, are substantially different from our Galaxy.

Our nearest massive neighbour is M31, the Great Galaxy in Andromeda. It is 4 times closer than NGC 55 and NGC 253, the next nearest sizeable galaxies. M31 is a spiral galaxy similar to our own. Since we view it from outside, the structure and distribution of stellar populations is clear. In contrast, our understanding of the arrangement of stars in the Milky Way is based on observations severely limited by obscuring clouds of dust. Observations of M31 should clarify the form of our Galaxy and the knowledge of our place within it.

On a clear, moonless autumn night, M31 can be seen with the naked eye as a patch of diffuse light just below the curved chain of stars which form the constellation Andromeda. A time exposure through a large telescope gives the result seen in Colour Plate 35. The eye sees only integrated light from stars in the central part of the Galactic Bulge, an ellipsoidal arrangement of old red stars. Photographs reveal a larger extent. Faint spiral arms extend to a diameter of 1.5° and contain bright patches and dark lanes. These are regions of recent star formation and clouds of absorbing dust. The spiral arms are blue, the colour of young, bright stars. Doppler shifts of spectral features can be measured accurately and show the SW arms to be moving towards us and the NE arms to be moving away. The rotational velocity is about 250 km s^{-1} , implying a period of revolution of 10^9 years, very similar to that of the Milky Way. At the other extreme, a short exposure through the telescope shows only a bright, almost starlike nucleus.

Table 13.1 lists properties of the Local Group, other nearby systems, and galaxies mentioned in this chapter. Dynamic masses are from several sources whose accuracy varies.

The Local Group galaxies have been extensively mapped, and a number of nearby bright galaxies have been observed by Einstein, ROSAT, ASCA, Chandra, and XMM. The X-ray appearance of these galaxies is quite different from the optical. Rather than a cloud of unresolved faint stars, each galaxy looks like a loose collection of individual bright sources, sometimes imbedded in diffuse emission. The deeper surveys of the Local Group find hundreds of sources in each galaxy. About one-third of

Table 13.1 Properties of some normal and starburst galaxies.

Name	NGC	Type	Distance (Mpc)	Diameter (arcmin)	Mass ^a galaxy (M_{\odot})	L_{opt} galaxy (L_{\odot})	L_X^b galaxy (erg s ⁻¹)	$L_X^{b,c}$ nucleus (erg s ⁻¹)
Milky Way	-	Sbc	-	-	6×10^{11}	1.4×10^{10}	3×10^{39}	3×10^{33}
LMC	-	SBm-Irr	0.050	645	2×10^{10}	1.3×10^9	7×10^{38}	No nucleus
SMC	-	Im	0.058	316	2×10^9		2×10^{38}	No nucleus
M31	224	Sb	0.77	190	8×10^{11}	3×10^{10}	3×10^{39}	3×10^{35}
M33	598	Sc,cd	0.84	60	3×10^{10}	4×10^9	2×10^{39}	1.5×10^{39} ?
	253	SABc	2.8	27	1×10^{11}	1×10^{10}	2×10^{39}	Starburst
M81	3031	SAab	3.6	27	4×10^{11}	2.4×10^{10}	4.6×10^{40}	4×10^{40}
M82	3034	I0	3.6	11	1.5×10^{10}	8×10^9	1×10^{40}	7×10^{39} , starburst
M83	5236	SBc	3.7	13	3×10^{11}	3×10^{10}	1×10^{40}	Starburst
Antennae	4038/39	Sc pec	19	4+	8×10^{10}		1×10^{41}	Starburst
	4278	E1	16	4	4×10^{11}	2×10^{10}	2×10^{40}	
M60	4649	E2	16	7	1.5×10^{12}	6×10^{10}	2×10^{41}	

^a Data are from several sources which use different criteria. Relative comparison accuracy $\sim \times 2$.

^b Energy range $\sim 0.5\text{--}5$ keV.

^c Often variable.

Table 13.2 Expected Chandra rates from X-ray sources in other galaxies.

Source	L_X^a (ergs s ⁻¹)	Milky Way		Rate/size (c ks ⁻¹)			
		Distance (kpc)	Rate/size (c s ⁻¹)				
XRB ^b	10^{38}	8	700/-	43000/-	180/-	9/-	
Crab PWN	3×10^{37}	2	4000/2'	9000/5''	39/0.3''	1.8/-	
Cyg Loop	2×10^{36}	0.8	2000/150'	510/2.4'	2.1/9''	0.10/1.9''	
Car Nebula	2×10^{35}	2.5	30/60'	80/3'	0.36/12''	0.02/2.4''	
O star	10^{33}	2.5	0.4/-	1.2/-	5×10^{-3} /-	3×10^{-4} /-	

^a Energy range 0.2–10 keV.

^b Assumed to be an Eddington-limited $1.4 M_{\odot}$ neutron star.

the sources, including most of the brighter ones, can be identified quickly by comparing positions with catalogues of optical, IR, and radio sources. X-ray spectra and extent also reveal the nature of objects.

13.2.1 Limits of source detection

Table 13.2 gives expected Chandra count rates and angular extent of representative Milky Way sources as they would appear in other galaxies.

The Milky Way column lists observed distance, count rate and extent. Rates for other galaxies have been calculated using Table 13.1 distances and a correction for the lesser absorption in the line of sight (which depends on the intrinsic spectrum and is usually uncertain by a factor of ~ 2). Note that expected rates are given in counts ks^{-1} . Thus a 100-ks Chandra observation of a Cygnus-Loop-type source in M31 would yield 200 counts – enough to recognise the 9'' extent of the remnant.

An image of a distant galaxy will have sources in the field which are not members of the galaxy being observed. These are *interlopers*, much closer than or much more distant than the galaxy. With a little extra work, most can be identified and excluded from further consideration. Foreground stars appear as prominent objects on optical plates, although M stars, which have high ratios of X-ray-to-optical luminosities, can be difficult. More distant sources are usually nuclei of active galaxies and quasars which can be identified through spectra of the optical counterparts. However, if viewed through the central region of a distant galaxy, obscuration by dust or confusion with the high density of stars can make an accurate identification difficult. A 100-ks Chandra observation of a high-galactic-latitude target typically has ≈ 100 AGN appearing as pointlike sources in the $16' \times 16'$ field, i.e. a spatial density of ≈ 0.4 AGN interlopers arcmin $^{-2}$.

The minimum flux detectable in a Chandra 100-ks observation, in a region free of diffuse emission, is $\approx 6 \times 10^{-16}$ ergs cm $^{-2}$ s $^{-1}$. This will yield a few counts, enough to positively detect the source since the background for point sources is negligible. This corresponds to a luminosity limit of 2×10^{32} ergs s $^{-1}$ for the Magellanic Clouds, 3×10^{34} ergs s $^{-1}$ for M31/M33, 2×10^{36} ergs s $^{-1}$ for a galaxy 5 Mpc distant and 2×10^{37} ergs s $^{-1}$ for galaxies in the Virgo cluster at 16 Mpc.

13.3 Surveys of the Magellanic Clouds

The Magellanic Clouds are near enough so that the luminous accretion-powered binaries and supernova remnants have been easily detected and spectra obtained in observations of moderate duration. Supernova remnants are easily identified by their spatial extent. A deep observation detects individual bright O stars. The Magellanic Clouds, however, are substantially different from our Galaxy, and their populations of X-ray sources have some significant differences from those in the Milky Way.

The Large Magellanic Cloud was first seen at X-ray wavelengths in October 1968 (Mark *et al.*, 1969)

using a rocket-borne detector. Although the X-ray luminosity and extent of the LMC were close to expectations, the result was accepted with reservation.

In September 1970, a larger detector was used to scan slowly across both the Large and Small Clouds, and X-ray emission above background was clearly detected (Price *et al.*, 1971). The result, however, seemed peculiar. Emission from the LMC appeared to come, not from many weak sources spread throughout the volume of the galaxy, but from a few bright, unresolved objects. The SMC appeared as a single bright source with a hard spectrum. At about the same time, in December 1970 and January 1971, the newly launched Uhuru satellite mapped the X-ray emission from both clouds. Data showed three bright sources in the LMC and, again, the single SMC source (Leong *et al.*, 1971).

These observations required that the individual sources have $L_X \approx 10^{38}$ ergs s $^{-1}$. This led to the realisation that these were kin to the bright sources in the central Milky Way, which, if clustered about the Galactic Centre, would have similar luminosity. These Galactic Bulge sources are now known to be accretion-powered binaries operating at about the Eddington limit. Subsequent observations with rockets and the satellites SAS-3, Ariel V and HEAO-1 found another bright source in the LMC and several bright transient sources in both clouds and gave accurate locations leading to identifications of optical counterparts. Two of these bright sources, LMC X-1 and LMC X-3, have since been identified as binary systems containing $\sim 10 M_\odot$ black holes.

To those who struggled to untangle source locations from the count rate data of early rocket observations, a glimpse at the 1980 Einstein maps of the Magellanic Clouds was like a look into a treasure chest, with precious gems and bright coins sparkling in the light (Figs. 13.2 and 13.3). It was difficult to decide which item to examine first. More than 120 sources were found within the LMC and over 40 within the SMC. A pleasing result was the large number of supernova remnants, easily identified by their spatial extent. Many were brighter X-ray emitters than expected. Usually, pre-observation predictions are optimistic, and fewer counts are collected than

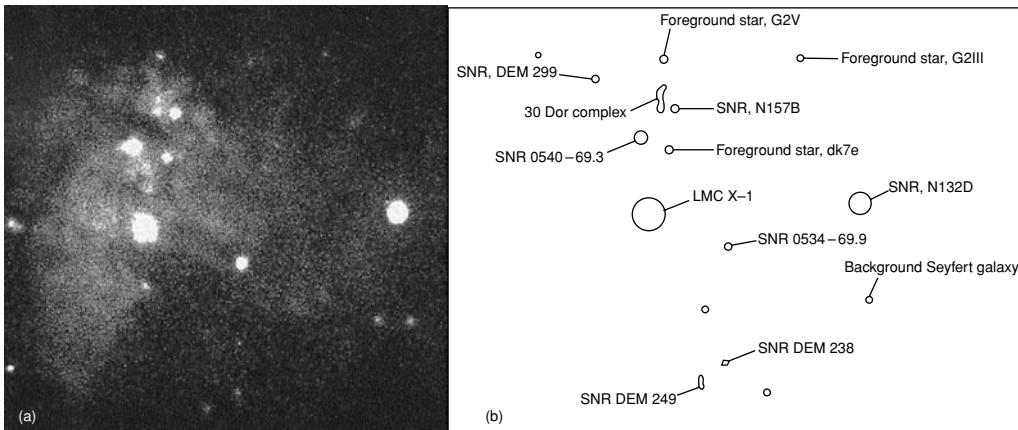


Fig I3.2 (left) A $2 \times 2^\circ$ field within the Large Magellanic Cloud as mapped with the relatively short exposure of the ROSAT PSPC survey. The brightest source, just left of centre, is the X-ray binary LMC X-1. Note the diffuse emission throughout the left side of the field. (right) Key giving identification of some sources. There are seven supernova remnants of varying brightness, three foreground stars and a background AGN (MPE/ROSAT 1991).

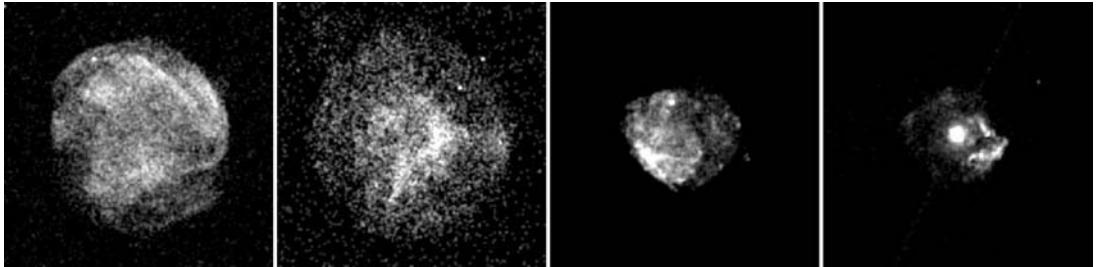


Fig I3.3 Magellanic Cloud SNRs imaged by the Chandra ACIS. Angular extent of each box is $3.1'$ (45 pc) and grey scale is a square-root stretch. Remnants are, from left to right, SNR 0534-69.9, thought to be the remnant of a type Ia SN; SNR 0049-73.6 or IKT6, the remnant of a type II SN; N49, containing a soft-gamma repeater (Chapter 9) which appears as a bright point source in the northern part of the remnant; and SNR 0540-69.3 (also in Colour Plate 25), with a very bright, overexposed central pulsar and PWN about as luminous as that within the Crab Nebula (from NASA/CXC, 2007).

anticipated. In this case the reverse was true, and detailed X-ray maps and spectra were obtained where none were expected. These remnants have since been observed by almost all subsequent missions.

A complete mapping of the Magellanic Clouds was accomplished during the 1990 ROSAT all-sky survey. Because the LMC is located close to the south ecliptic pole, it was scanned more often than equatorial locations, and average exposure was ≈ 2000 s. ROSAT catalogues of Magellanic Cloud sources incorporate these data as well as data from 10–20 ks pointed observations (Kahabka *et al.*, 1999; Haberl *et al.*, 2000; Habrol & Rietsch,

1999). A 10-ks ROSAT PSPC pointing collected ≈ 10 counts from a 2×10^{34} ergs s^{-1} source in the LMC. This is roughly the limit of the ROSAT Magellanic Cloud surveys, but coverage was not uniform.

The ROSAT surveys of the LMC/SMC detected $\approx 750/500$ sources in these fields. About one-fourth of these were identified via optical counterparts, spectra and extent. Thirty per cent of the identified sources are foreground stars, and 10 per cent are background AGN. Initial identifications included 54/19 SNR, 17/13 X-ray binaries and 12/3 supersoft sources. There are many HMXBs ($\approx 16/13$) and only one LMXB, the Uhuru-discovered LMC X-2. This is in marked contrast to the Milky

Way population of $\approx 50:160$ known HMXB:LMXB sources. The Magellanic Cloud population is rich in Be star binaries and poor in the low-mass systems. The X-ray luminosity of both Magellanic Clouds depends on the state of a few strong, highly variable sources. Total luminosity is usually $\sim 7 \times 10^{38}/2 \times 10^{38}$ ergs s $^{-1}$. The diffuse emission is steady and $3 \times 10^{38}/1 \times 10^{37}$ ergs s $^{-1}$. Diffuse emission is brightest, and hardest, around the star-forming region 30 Doradus in the LMC, as illustrated in Fig. 13.2.

Magellanic Cloud supernova remnants are particularly attractive targets. The diameters of 0.5'-4' fit nicely in the high-resolution area of an X-ray telescope, and the remnants are luminous. Thousands of counts can be collected in relatively short times. A few remnants are bright enough to yield grating spectra, and one is shown in Fig. 8.20. Chandra has observed ~ 40 remnants in the Magellanic Clouds with exposure times of 20–70 ks (e.g. Williams *et al.*, 2005); ~ 10 are thought to be from type Ia SN and ~ 10 from type II. Compact objects and/or PWNe have been imaged in eight of the type II remnants. Figure 13.3 shows some of these Chandra observations.

XMM has observed 13 SMC remnants with 15–30 ks exposures. On the basis of X-ray spectra and morphology, three of these are classified as remnants of type Ia SNe (van der Hayden *et al.*, 2004). Key features in the spectra are O and Fe emission lines at 0.6 and 0.9 keV. Interstellar absorption renders these lines invisible for most Milky Way remnants.

A Chandra survey of the central part of the SMC used five 10-ks ACIS-I fields to cover 0.36 deg 2 . In this small area, 158 sources were detected down to a limit of 4×10^{33} ergs s $^{-1}$, ~ 20 times fainter than the ROSAT survey limit (Antoniou *et al.*, 2009), and 113 were associated with optical counterparts. In this sample, there were 30 HMXBs with Be star companions, which led to the conclusion that these are twice as common with respect to early stars in the SMC than in the Milky Way.

13.4 M31

X-rays from M31 were first observed in February 1973 with a rocket-borne detector (Bowyer *et al.*,

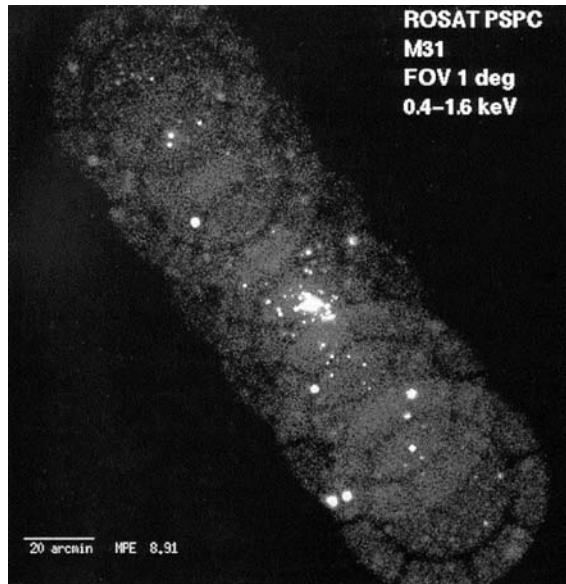


Fig 13.4 A ROSAT map of M31. This mosaic of six overlapping 30-ks PSPC exposures covers most of M31. Fields of view have been limited to 1°, and the grid pattern is due to shadowing by the window-support structure. There is bright extended emission from the nuclear region, and there are only a few bright sources in the outer spiral arms (MPE/ROSAT 1992).

1974). This was subsequently confirmed by the Uhuru and Ariel V all-sky surveys. The calculated X-ray luminosity was 2×10^{39} ergs s $^{-1}$, about the same as that observed from the whole of the Milky Way. In contrast to the first LMC observations, this was the integrated emission from all the bright sources in M31.

The Einstein IPC survey in 1979 revealed 50 bright X-ray sources in the spiral arms and an unresolved cluster of sources in the nuclear region. The Einstein HRI resolved this cluster into ≈ 20 bright individual sources (van Speybroeck *et al.*, 1979).

Although the total luminosity is about as expected, the bright source distribution is not. The dominant central cluster is associated with the Bulge. The bright sources are thought to be population II (old stars) accretion-powered binaries, the same as the bright Bulge sources in the Milky Way. There are, however, more than expected, and they are grouped closer to the nucleus.

This is clearly seen in the ROSAT PSPC survey result shown in Fig. 13.4. M31 was imaged in six

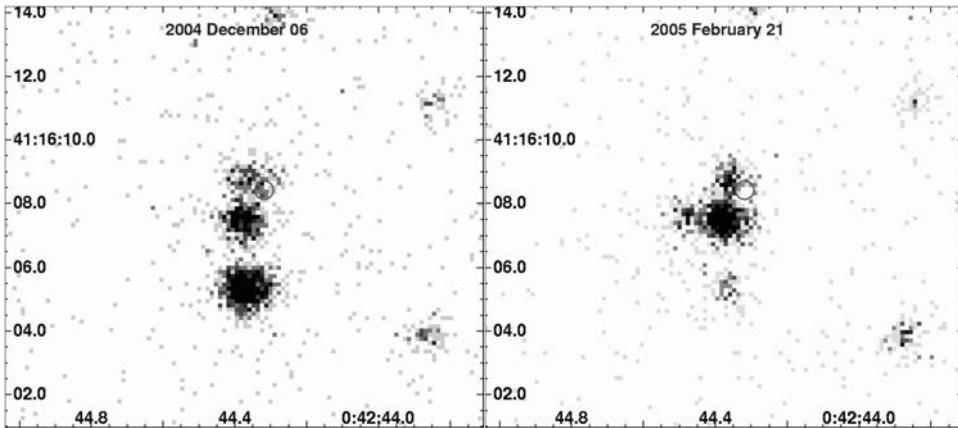


Fig 13.5 Two 50-ks Chandra HRC observations of the heart of M31. The small circles show the location of the nucleus, M31*. The bright central source in this figure is not the nucleus but a supersoft source. There is also a fading transient 2'' south of the nucleus (courtesy of M. Garcia).

evenly spaced pointings, and total exposure was ≈ 200 ks (Supper *et al.*, 1997). The fields of view included the entire galaxy. Three hundred ninety-six sources were detected, of which ≈ 80 are probably background sources unconnected with M31.

XMM-Newton surveyed M31 with 10 observations covering two thirds of the galaxy and totaling 350 ks (Pietsch *et al.*, 2005). Eight hundred ninety-six sources were found with $L_X \geq 5 \times 10^{34}$ ergs s $^{-1}$, of which 133 are identified as probably background AGN or foreground stars. There are 567 hard sources which are either XRBs, pulsar-wind nebulae or yet more background AGN. The brightest source had $L_X = 2.8 \times 10^{38}$ ergs s $^{-1}$ (in the band 0.2–4.5 keV).

Chandra observed the central $17' \times 17'$ with eight ACIS-I observations spread over a 1.5-year time span (Kong *et al.*, 2002). Total exposure was 40 ks, which resulted in the detection of 204 sources with $L_X \geq 2 \times 10^{35}$ ergs s $^{-1}$. Although the area of the Chandra survey was only 1/15 that of the XMM survey, and with a higher threshold, Chandra found a remarkable 1/4 the number of sources. This was because the density of sources in the centre of M31 is much higher than in the spiral arms and a perfect fit with Chandra's higher resolution. The brightest source in the band 0.3–7 keV had $L_X = 1.4 \times 10^{38}$ ergs s $^{-1}$ and approximately half of all Chandra sources were variable

from month to month. At any one time, approximately 3 transients were visible, and a total of 13 transients were found. Several SNR were detected and spatially resolved. The observations were not long, and not many photons were detected, so the size of the remnants was the only significant information derived from the images. Spectra, however, obtained from longer off-axis pointings and from XMM-Newton were detailed enough to be useful.

The optical nucleus of M31 is diffuse and double, with peak emission regions separated by just $1.1''$. At the centre, there is an unresolved radio source, M31*, thought to be a $3 \times 10^7 M_\odot$ black hole. Chandra HRC pointings (Garcia *et al.*, 2005) have detected very weak emission from M31*, indicating a luminosity of $L_X = 3 \times 10^{35}$ ergs s $^{-1}$. Figure 13.5 shows how the location of M31* was determined with an accuracy of $0.1''$ by using other detected sources to register the images. In particular, a bright supersoft source $1''$ south of M31* once thought to be the nucleus is not. Like the supermassive black hole in the Milky Way, the nucleus of M31 is underluminous. It is not currently an AGN (but almost certainly was in the past). The favoured explanation for this is a radiatively inefficient accretion flow, similar to that applicable to SgrA* in the centre of the Milky Way.

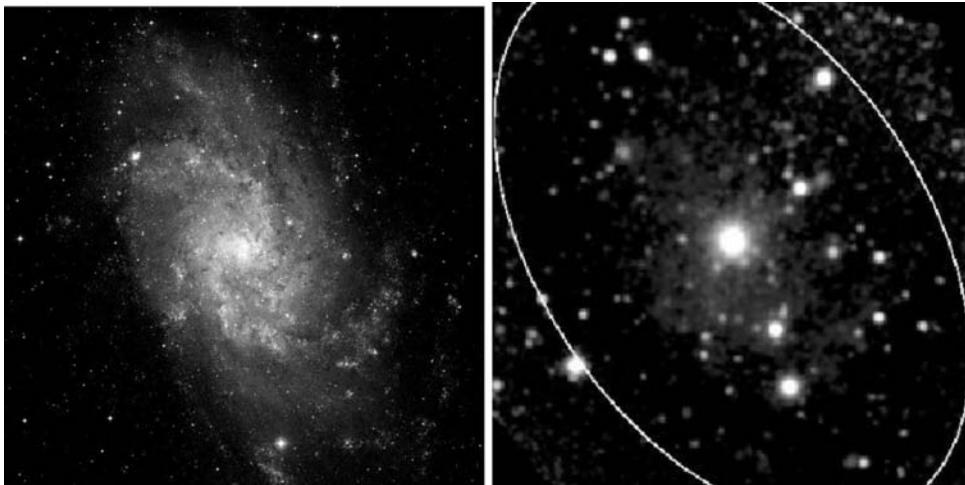


Fig 13.6 (left) The Local Group galaxy M33 from the Digital Sky Survey blue band. The field is 40' square. (right) An XMM X-ray image of the same field. Solid line shows the outermost limit of the optical galaxy. The eclipsing black hole is the first bright source SSW (lower right) of the central brightest source. This EPIC detector image was made by combining 15 10-ks pointings (STScI/DSS 2007; Pietsch et al., 2004).

13.5 | M33

Since M33 has a diameter of $\approx 60'$, comparable to an X-ray telescope field of view, modern surveys of M33 are complete down to a low flux level. This is not the case for M31 and the Magellanic Clouds, which are so large that they are only partially observed to a deep level, particularly with Chandra. An Einstein survey in 1980 found 17 point-like sources in the M33 field. The brightest was close to the nucleus, and four were background objects. In the interval 1991–1998, ROSAT found 184 sources. More recently, a 150-ks XMM image was obtained, which is compared with the optical galaxy in Fig. 13.6. This XMM survey found 408 sources with $L_X \geq 1 \times 10^{35}$ ergs s $^{-1}$ in a 0.8-square-degree field. A 1.4-Ms Chandra survey was accomplished in 2005–2006. An analysis of the data finds ≈ 600 sources with $L_X \geq 2 \times 10^{34}$ ergs s $^{-1}$ in a 0.25-square-degree field.

The brighter sources yield enough counts for spatial and spectral analysis. Comparison with optical and radio images is easily made. Figure 13.7 shows a supernova remnant well resolved by Chandra. To be fair, this remnant is the brightest remnant in M33 ($L_X = 1.7 \times 10^{37}$ ergs s $^{-1}$), and

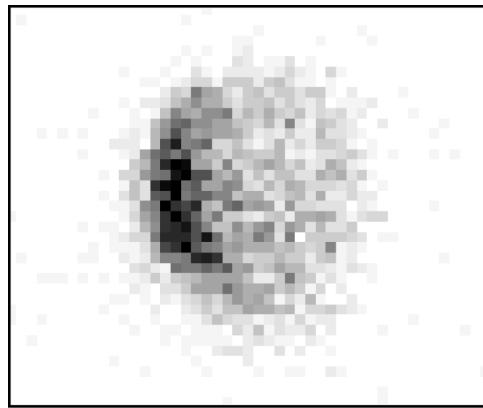


Fig 13.7 An 190-ks Chandra observation of a 5'' (20 pc) diameter supernova remnant in M33 which is clearly resolved. Each pixel in this unsmoothed image corresponds to 0.24'' (from Gaetz et al., 2007).

this was a deep observation, so the result is better than ‘typical’. Another example is the source M33 X-7, which has been identified as an eclipsing black-hole binary. The orbital inclination is small, and the known distance allows an accurate determination of spectral type and size of the companion star. The black-hole mass is thus accurately determined as $15.6 \pm 1.4 M_\odot$ (Orosz et al., 2007). At

Table 13.3 Some X-ray surveys of Local Group galaxies.

Galaxy	Spacecraft	Area (deg ²)	L_x Limit (ergs s ⁻¹)	Total sources	XRB (LM/HM)	SNR/SSS	Interlopers	Sources (deg ⁻²)
Milky Way					162/50	~100/2		
LMC	ROSAT	58		750	1/16	54/7		~13
SMC	ROSAT	9		248	0/13	19/5		~50
	Chandra	0.36	4×10^{33}	158	0/30			~20
M31	ROSAT	6.3	3×10^{35}	396	29/-	17/	~160	~60
	XMM	1.24	5×10^{34}	856	37/~16	44/18	~130	~700
	Chandra	0.08	2×10^{35}	204	22/-	2/9	~30	~2500
M33	ROSAT	2.2	2×10^{35}	184	[4]	17/7	~16	~80
	XMM	0.8	1×10^{35}	408	44/5	2/	~50	~500
	Chandra	0.25	2×10^{34}	662		82/	~300	~2600

Note. Square brackets indicate classification as XRB; no distinction between LMXB and HMXB.

Note. Sources deg⁻² include background sources.



Fig 13.8 Two starburst galaxies. Both images are from the Digital Sky Survey blue band. (left) M82, an edge-on galaxy with intense starburst activity. The field is 7.6' square. (right) M83, a larger face-on spiral with moderate activity around the centre. The field is 12.8' square. Corresponding X-ray images are shown in Colour Plate 37 and Colour Plate 36 (STScI/DSS, 2007).

this time this is the only eclipsing and the heaviest stellar-mass black hole known. The brightest source in M33 is the unidentified central source. It could be another bright binary, or it could be the nuclear black hole. The source is variable, and the light curve has been searched for periodicity with no clear result.

Thus the principal members of the Local Group have been surveyed and hundreds of sources catalogued. Radio and optical information is available or can be obtained to aid in source identi-

fication. Eventually, the source populations will be well determined. For more distant systems, source identification is more difficult. Table 13.3 lists results from Local Group surveys.

13.6 | Starburst galaxies

13.6.1 M82

M82, shown at the left in Fig. 13.8, is a rather small galaxy with intense activity at the centre.

It contains a large population of supernova remnants and luminous HII regions. Energy is supplied by supernovae at a rate of approximately one every 10 years, and the far infrared luminosity is greater than the optical. Since these signs are characteristic of star-forming regions, M82 is called a ‘starburst’ galaxy. The galaxy is viewed edge-on, and outflowing ionised material forms a strong wind directed perpendicular to the plane of the galaxy.

XMM-Newton and Chandra observations of M82 illustrate the strong points of each observatory. The large area of XMM makes it more sensitive to diffuse emission, whereas the arcsecond resolution of Chandra is better for the detection of point sources in crowded regions. The left half of Colour Plate 37 shows the diffuse emission imaged in a 30-ks XMM exposure. It is elongated perpendicular to the plane of the galaxy, as seen in Fig. 13.8. Sources in the centre are overexposed. This diffuse emission comes from a strong wind of ionised material. The colour in Colour Plate 37 shows that the temperature is hot at the centre but does not change much in the outer regions. X-ray and H α structure are correlated, e.g. the northern wind terminates at a feature called the H α cap, perhaps an indication of material in a halo. Note that the wind extends 20' NW-SE, which is more than the 9' NE-SW diameter of the optical galaxy.

Colour Plate 37 also shows the Chandra image of the bright centre. Chandra finds 25 point sources with $L_X \geq 2 \times 10^{37}$ ergs s $^{-1}$ associated with M82. The threshold is high because of the strong diffuse emission. Eleven sources clustered close to the centre have $L_X \geq 10^{38}$ ergs s $^{-1}$. These are believed to be HMXBs with neutron-star and black-hole secondaries, and the unusually high number is another sign of intense star-forming activity. The brightest of these is highly variable and has been detected radiating 9×10^{40} ergs s $^{-1}$ in the 0.2–10 keV band – the most luminous binary so far observed – an ultra-luminous X-ray binary or ULX. It is located 9'' from the kinematic centre, so it is not a supermassive black hole or AGN. This source alone was probably responsible for the variable intensity observed earlier by ROSAT and ASCA. Note that the source detection limit is high due to the high local diffuse background.

The XMM spectrometer obtained a high resolution spectrum of the diffuse wind, which is shown in Fig. 13.9. The spectrum contains hydrogen-like lines from N, O, Ne, Mg and Si. Fe L lines from Fe XVII to Fe XIV are also prominent. Emission is from a hot gas. Lines present indicate a range of temperatures from 0.3 to 1.5 keV and that the gas is in collisional ionisation equilibrium. Abundances are near-solar, except for Mg and Si, which are higher. L_X of the wind is 5×10^{40} ergs s $^{-1}$ – it is called a ‘superwind’. For good reason, M82 is considered the prototypical starburst galaxy!

13.6.2 M83

M83 is a barred spiral galaxy with a starburst centre. It is viewed face-on, so source confusion and absorption in the galaxy are at a minimum. It just fills the field of the Chandra ACIS detector and so has been surveyed with a single pointing and is close enough so spectra of the brighter sources were obtained. Figure 13.8 shows the 10' diameter optical galaxy. Colour Plate 36 shows the result of a 50-ks observation with the 8' Chandra ACIS-S detector. Eighty-one point sources were detected and located with 0.5'' accuracy. The detection threshold was 3×10^{36} ergs s $^{-1}$. About 12 sources are expected to be background AGN. Figure 13.10 shows spectra of some of the brighter (a few hundred counts) sources of different types. Such spectra can be used for classification.

There are 15 sources within 1' of the nucleus. Density in this inner region is ≈ 10 times higher than throughout the rest of the galaxy, and the relative number of high-luminosity sources is greater. The seven brightest sources have $L_X \geq 2 \times 10^{38}$ ergs s $^{-1}$. The brightest one has $L_X = 6 \times 10^{38}$ ergs s $^{-1}$. One of two apparent nuclei is an X-ray and IR source with $L_X = 3 \times 10^{38}$ ergs s $^{-1}$, well below the expected luminosity of an AGN-type supermassive black hole.

There is diffuse emission associated with spiral arms and concentrated at the centre. Fifty per cent of the central emission is diffuse and has an L_X of 1.2×10^{39} ergs s $^{-1}$. The diffuse spectrum is thermal with $kT \approx 0.6$ keV, as expected for hot gas associated with HII regions. Table 13.4 lists results from starburst galaxy surveys.

Fig 13.9 Spectrum of M82 from the XMM-Newton reflection-grating spectrometer. Bright emission lines are identified, and vertical lines show RIF (resonance, intercombination and forbidden) triplets of He-like species. Emission is from a hot gas characterised by a large temperature range (from Read & Stevens, 2002).

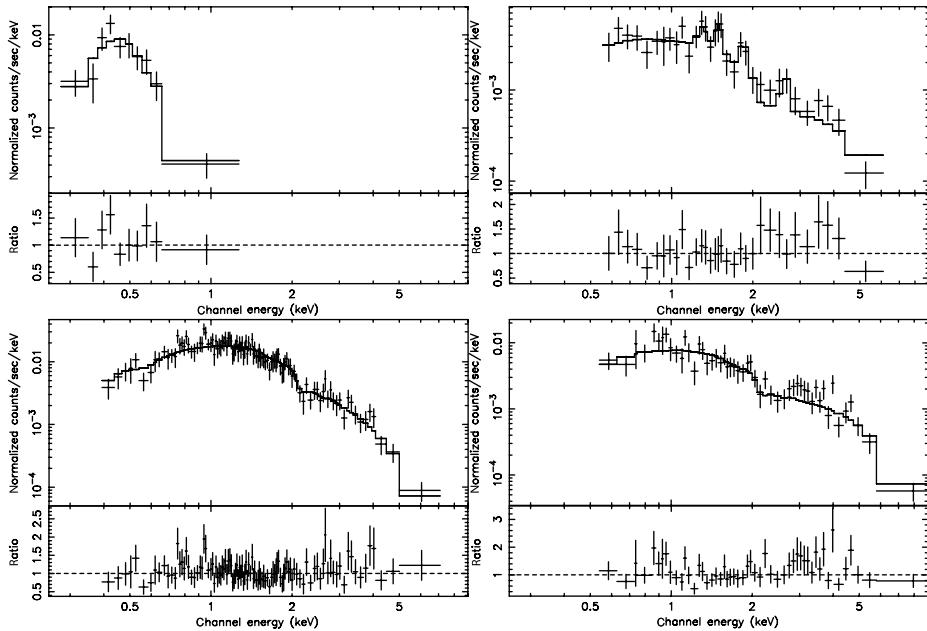
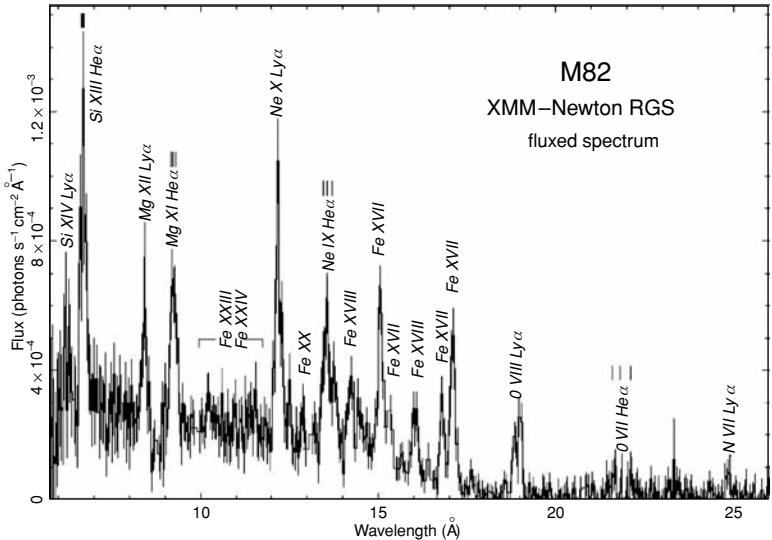


Fig 13.10 X-ray spectra from a Chandra survey of the starburst galaxy M83. Such spectra indicate type of source and are used to classify point sources. (top left) A supersoft source. There is no emission above 1 keV. (top right) A thermal source, probably a supernova remnant. The spectrum contains several emission lines. (bottom left) Probably an X-ray binary. The spectrum is a featureless continuum. (bottom right) The galactic nucleus, a featureless power-law continuum (from Soria & Wu, 2003).

13.7 The centre of the Milky Way

A discussion of the Local Group is not complete without mention of our own Galaxy, named for the band of faint stars seen spanning the sky on

a clear night. Recognising structure is difficult because of our location within a disc of gas and dust which obscures distant objects, particularly the centre. Nevertheless, a great deal of effort has gone into attempts to learn about the nucleus of the Milky Way.

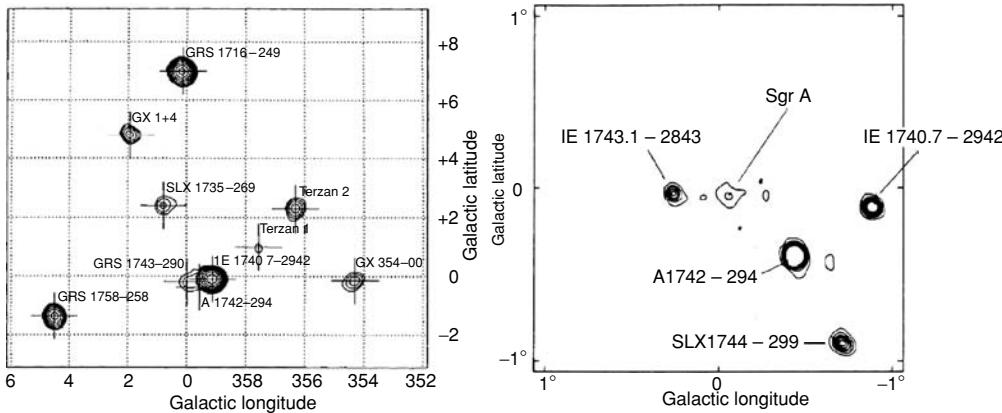


Fig 13.11 Two high-energy images of the centre of the Milky Way from coded-mask detectors. (left) A $12^\circ \times 14^\circ$ region in the range 40–75 keV from the Sigma instrument on GRANAT. Total observing time is 9×10^6 s (Vargas et al., 1996). (right) A 2° square field in the somewhat softer range 3–35 keV, as mapped by an instrument on Spacelab. The galactic centre (SgrA) barely appears (Skinner et al., 1987). All bright sources vary strongly, only one is common to the two images.

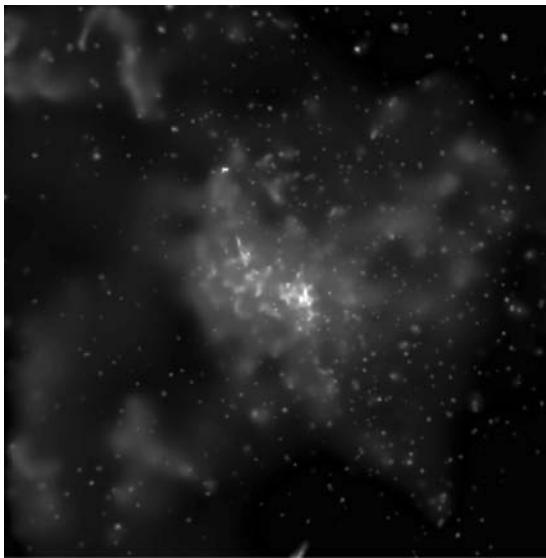


Fig 13.12 A 590-ks Chandra observation of the centre of the Milky Way. The energy band is 0.5–8 keV. This field is $8.5'$ square, and the Galactic Centre, SgrA*, is the bright region at centre. The image has been adaptively smoothed, and the grey scale is logarithmic (from Muno et al., 2003).

The centre is complex. Radio observations with resolution of a few arcminutes found a strong source, which was named SgrA. At arcsecond resolution, a weak pointlike non-thermal radio source was detected, and this is called SgrA*. By measuring the motions of surrounding stars, the object at

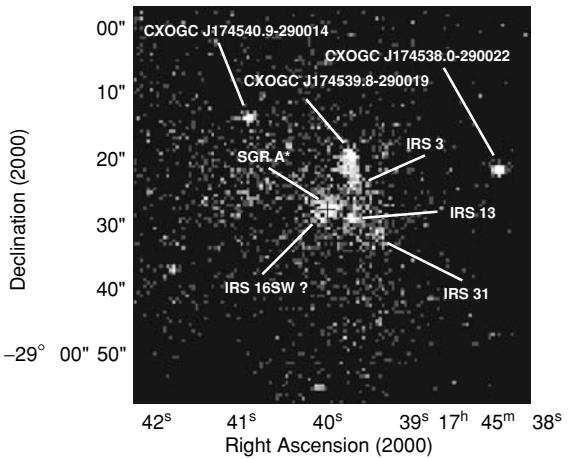


Fig 13.13 A 51-ks Chandra image of a $1'$ field centred on the radio source SgrA*, whose position is marked by a black cross. There is a weak X-ray source close to this location. Each pixel is $0.49''$ square, and data have not been smoothed. The energy band is 0.5–7 keV. Other sources (but not SgrA*) are indicated by arrows. Those labelled CXOGC were new. Others correspond to infrared sources (from Baganoff et al., 2003).

the centre has been determined to be a $3 \times 10^6 M_\odot$ black hole. There is too much intervening dust to get visible or soft X-ray information. Above 2 keV, however, the centre is observable and has been studied by several instruments. Figure 13.11 is a map of the central region from a Spacelab

Table 13.4 | Some X-ray surveys of starburst galaxies.

Galaxy	Spacecraft	Area (deg ²)	L_x limit (ergs s ⁻¹)	Total sources	Interlopers	Sources (deg ⁻²)
M82	ROSAT	0.090	6×10^{37}	13	1	~140
	Chandra	0.0070	2×10^{37}	25		~3600
M83	ROSAT	0.034	1.4×10^{37}	14	1	~400
	Chandra	0.071	3×10^{36}	127	~12	~1800
Antennae	XMM	0.020		~13		~700
	Chandra	0.005	5×10^{37}	38	~2	~7600

coded-mask telescope. In this 3–35 keV band, SgrA, a hard source at the centre is clearly detected and resolved within the $\approx 5'$ capability of the instrument. A similar detector on the Soviet spacecraft GRANAT also resolved SgrA and surrounding sources, including a nearby transient. The Einstein IPC detector, with reasonable sensitivity at 2–6 keV and 1' resolution, mapped the Galactic Centre in 1980 and detected emission from the vicinity of SgrA*. There is a lot happening in this region, however, and arcsecond resolution is necessary to separate the galactic nucleus from other sources.

Chandra observations show how very complex this region is. For the first time SgrA* was resolved in X-rays (Muno *et al.*, 2003, Baganoff *et al.*, 2001). Figure 13.12 shows a deep observation of the centre. The 17' square field, which surrounds the Fig. 13.12 field, contains 2357 pointlike sources and diffuse emission. There are 281 foreground sources and ~ 100 background AGN. Galactic sources lie in a column stretching through the centre to the far side of the Galaxy. Many of the sources in this field are accreting white dwarfs spread along the direction of view. The detection limit at the centre is $L_X \approx 10^{30}$ ergs s⁻¹.

Figure 13.13 shows the central arcminute of this field. There is a weak source located $0.27 \pm 0.18''$ from the radio location of SgrA*. On the basis of this location, the X-ray source is identified with SgrA*. In the 2–10 keV band, L_X is a mere 3×10^{33} ergs s⁻¹ seen through an absorbing column of $N_H = 1 \times 10^{23}$ cm⁻². The source is extended by $\approx 1.4''$, which is thought to indicate the accretion radius of the black hole and corresponds to $\approx 10^4$ AU at the distance of the Galactic Centre. During

subsequent Chandra observations, SgrA* has been observed to flare several times. During a flare, X-ray emission increases for about 10^4 s and at peak is 10–20 times the quiescent level – not unusual behaviour for a black hole.

The field is full of diffuse emission. The sources seen in Fig. 13.13 are embedded in a $1' \times 2'$ bright region which is the radio source SgrA East. This is believed to be a supernova remnant that actually encompasses SgrA* and is prominent at the centre of Fig. 13.12. It is also easily visible in Colour Plate 38. X-ray spectra of diffuse gas in the central region contain characteristic K α lines of Fe. This has been interpreted as a sign of past nuclear activity. If SgrA* were very bright in the past, fluorescent radiation would be expected from surrounding material. It is more likely, however, that other sources are responsible.

13.8 | Elliptical galaxies

Elliptical galaxies are old, and the X-ray binaries therein are almost all LMXBs. The HMXBs, which have a young massive-star component, have short lives and should be almost all gone. Accordingly, an observation of an elliptical galaxy should yield a ‘clean’ sample of LMXB sources. Figure 13.14 shows a Chandra observation of NGC 4278. The optical galaxy is regular, with almost no structure. There is probably a black hole at the centre, but it is not active now. The distance is 16 Mpc, so only the bright X-ray binaries were seen. Two ACIS-S fields were merged and partially appear in this figure. Only sources inside the outlined galaxy

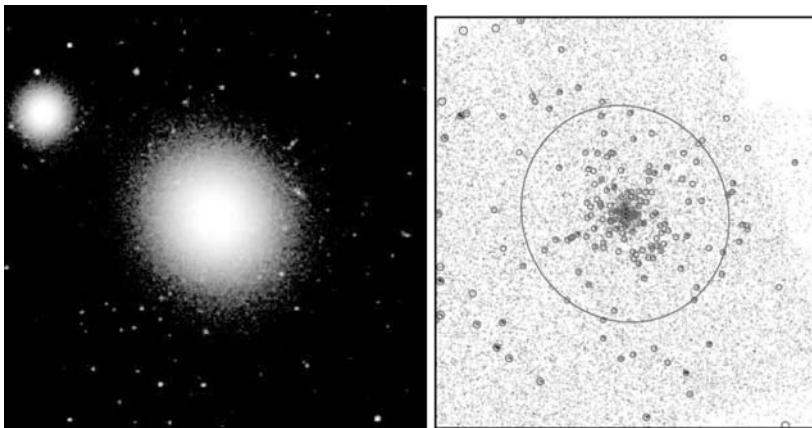


Fig I3.14 Optical and X-ray images of the elliptical galaxy NGC 4278. (left) Red DSS field $7.5' \times 7.5'$. The large galaxy is NGC 4278. The smaller galaxy to the NE, NGC 4283, does not have a concentration of X-ray sources. (right) 145-ks Chandra X-ray image to same scale. The large ellipse shows the D_{25} isophot of the optical galaxy. X-ray sources, which are difficult to see, are marked with small circles (Kim et al., 2006).

were considered. Although some galactic binaries may be outside this contour, the fraction of background AGNs is high there. Sources within $10''$ of the centre were also excluded because of source confusion. Out of 197 sources above the threshold of $5 \times 10^{36} \text{ ergs s}^{-1}$ in the ACIS-S fields, 112 were in the allowed region and were used to construct the luminosity function shown in Fig. 13.15. The estimated number of background AGN is ≈ 8 . Diffuse emission from NGC 4278 is small compared to that from sources.

NCG 4649 (M60), on the other hand, is more massive, and diffuse emission dominates, as shown in Fig. 13.16. NCG 4649 is in the Virgo cluster and fairly large. One hundred sixty-five pointlike sources were detected. The diffuse emission fills most of the galaxy and is bright at the centre. No pointlike component is seen there, so the central black hole was not detected. There is faint structure within the diffuse component, which may be evidence of an active nucleus in the past. Other elliptical galaxies show varied, more pronounced diffuse structure. Assuming the emitting gas is bound within the galaxy by gravity, the mass of the galaxy can be calculated. The method is the same as for galaxy clusters and is discussed in Chapter 15. The derived mass is usually greater than that of the stars, implying much dark matter.

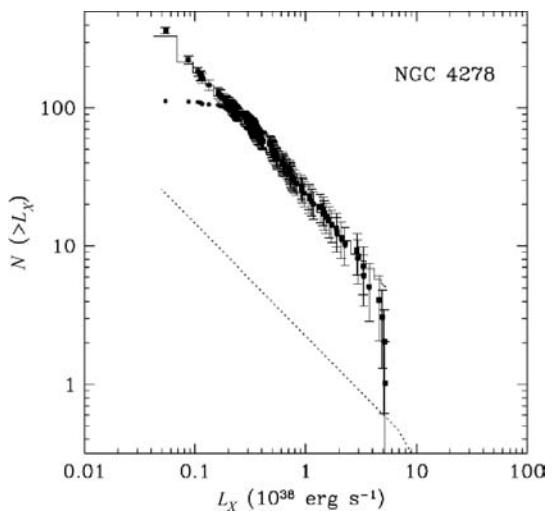


Fig I3.15 X-ray luminosity function for sources in the elliptical galaxy NGC 4278. All sources are assumed to be LMXBs. The faint dashed line shows the estimated background AGN contribution. Data points at low L_X show a completeness correction applied to account for low detection efficiency of faint sources. Histogram is a power law fit with slope -0.9 (Kim et al., 2006).

13.9 Colliding galaxies

When galaxies collide, the stars do not. They merely pass through the volume of the other

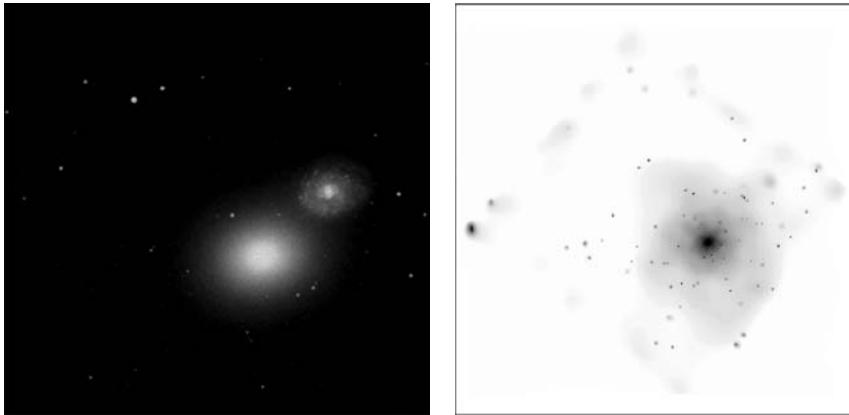


Fig 13.16 Optical and X-ray images of the elliptical galaxy NGC 4649. (left) DSS red field $12' \times 12'$. The large galaxy is NGC 4649; the smaller is NGC 4647. (right) A 19-ks adaptively smoothed Chandra X-ray image to the same scale. The 8' Chandra ACIS-S field is oriented diagonally within the box, and the apparent faint emission at the edge of the detector is an artefact. Note the extensive diffuse emission which is brightest at the centre of the galaxy. The LMXB sources appear as faint dots (Randall *et al.*, 2004).

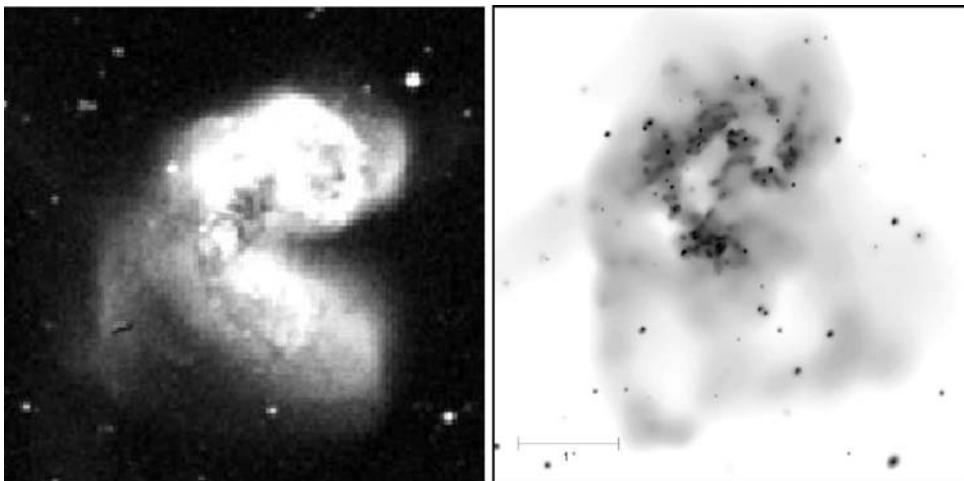


Fig 13.17 Optical and X-ray images of the colliding galaxies NGC 4038/NGC 4039, the ‘antennae’. (left) From the Digital Sky Survey blue band. The field is $4.5'$ square. The antennae themselves extend $8'$ from the centre and are too faint to appear here. (right) A 420-ks (5 day) Chandra ACIS-S observation of this field. The data have been adaptively smoothed. There is strong diffuse emission, and several of the pointlike sources are exceptionally luminous (Fabbiano *et al.*, 2001).

system. The extended and diffuse interstellar gas clouds, however, do collide, and the resulting shock waves compress gas and trigger star formation. Thus colliding galaxies usually have strong indications of star-forming activity and large numbers of young massive stars. Such is the case

for the galaxy pair NGC 4038/4039, called the ‘antennae’ because of two faint trailing arcs of emission extending from the merging centres of the galaxies. Figure 13.17 shows the DSS appearance of the centres and a deep Chandra observation of the same region (Fabbiano *et al.*, 2001).

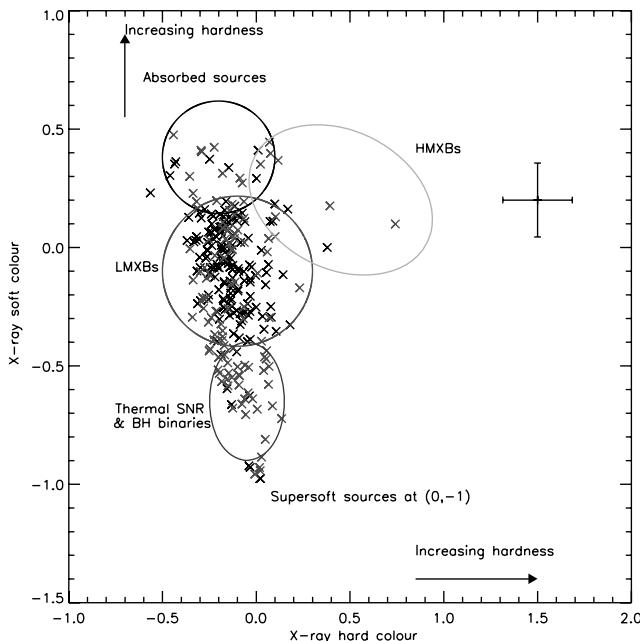


Fig 13.18 A colour-colour plot for point sources detected in the Chandra observations of two spiral and three elliptical/bulge galaxies. Circles show areas where different types of sources dominate. The cross shows the uncertainty for a typical data point (Prestwich et al., 2003).

The Chandra image contains 48 apparent sources; 10 are extended and associated with bright diffuse emission and 38 are pointlike. The distance is 19 Mpc. The luminosity threshold for detection is 5×10^{37} ergs s $^{-1}$, so only bright XRBs should be seen. This is indeed the case. In fact, many of these sources are extraordinarily luminous. There are eight objects with $10^{39} \leq L_X \leq 10^{40}$ ergs s $^{-1}$. The spectra are hard, typical of XRB. A 2-year monitoring program measured variability as expected of XRBs but not the decay expected of young SNR. These are apparently all ULXs, rare in most galaxies but in abundance here.

Diffuse emission is also bright. The spectrum is thermal with emission lines from Mg Si, and Fe and indicates a hot ($kT \sim 5$ keV) ISM. There is some indication that abundances vary from place to place. Diffuse X-rays follow the same general pattern as the H α emission, but not exactly. There are holes in the H α brightness which coincide with peaks of X-ray emission. These are perhaps super-bubbles but an order of magnitude more energetic than those in the Magellanic Clouds. There is an $\approx 5'$ north-south extent of faint diffuse emission which indicates an outflow of material. Luminosity of the diffuse emission is 5×10^{40} ergs s $^{-1}$, about the same as the sum of the pointlike sources.

13.10 | Source populations

13.10.1 Identification via spectra

Figure 13.10 shows spectra typical of several source types. If enough counts are collected, such spectra can reveal the nature of the source. For fainter sources, hardness ratios are used in which events are sorted into three energy bands: soft (S), medium (M) and hard (H). Ranges are typically 0.3–1, 1–2 and 2–8 keV, respectively. Ratios used vary, but in the example shown here are $(M - S)/(S + M + H)$, or soft colour, and $(H - M)/(S + M + H)$, or hard colour.

Figure 13.18 shows hardness ratio data as a colour-colour plot. The very softest sources are SSS (see Chapter 10). Moderately soft spectra generally indicate SNRs. Spectra of X-ray binaries depend on the amount of absorption and state of the source. LMXBs fall in the centre of the diagram. If an HMXB forms an accretion disc, the spectrum will be soft. If accretion is from a wind, the spectrum is hard and the source falls at the upper right of this diagram. As absorption increases, the soft colour index increases, and points migrate upwards. At about 10^{22} atoms cm $^{-2}$, the hard colour index also increases, and points will move to the right. Areas

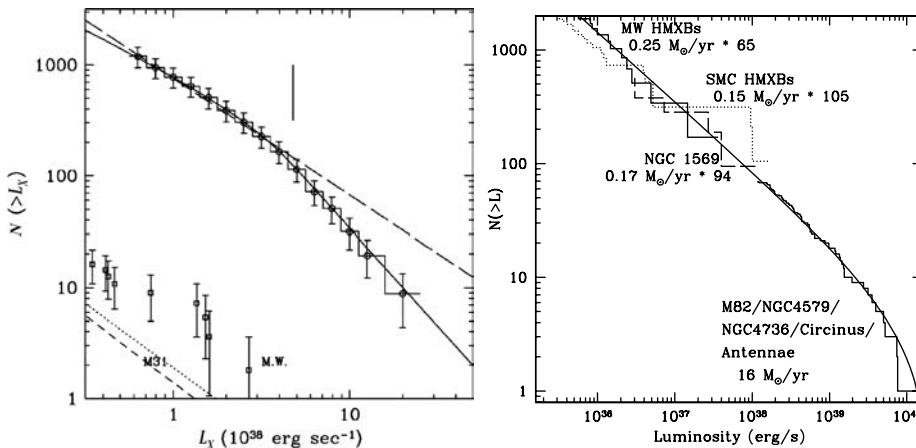


Fig 13.19 X-ray luminosity functions for accretion powered binaries. (left) Cumulative luminosity function from 14 E and S0 galaxies (Kim & Fabbiano, 2004). XLFs for Milky Way and M31 are in lower left corner. (right) Luminosity functions from eight starburst galaxies (Grimm *et al.*, 2003). The starburst XLF is flatter than and extends to greater luminosity than the LMXB XLF.

where a particular type of source usually falls are indicated by circles. Absorption, however, will move sources from the expected regions.

13.10.2 X-ray binaries

X-ray luminosities have now been measured for hundreds of extragalactic binaries. In the Local Group, identifications have been verified by observation of optical counterparts. For faint, distant sources, it is assumed that strong sources in globular clusters, elliptical galaxies and galactic bulges are LMXBs. Sources in galactic discs and starburst galaxies, regions with young massive stars, are assumed to be dominated by HMXBs. These populations need not be exclusively HMXB or LMXB, but the assumption holds for most of the bright sources.

Figures 13.15 and 13.19 show representative luminosity functions or XLFs. It is customary to plot the integral source distribution. In this representation, if all sources were to have a given luminosity, the curve would be flat up to that luminosity, then fall to zero. If sources were evenly distributed in logarithmic luminosity increments, the curve would flatten at lower luminosity. A power law with slope -1 implies a 10-fold increase in number of sources for every factor of 10 decrease in luminosity.

Figure 13.15 is the LMXB luminosity function for the elliptical galaxy shown in Fig. 13.14.

This distribution is well fit by a power law of slope -0.9 ± 0.1 . There are many more faint than bright LMXBs. The most luminous source has $L_x = 5 \times 10^{38}$ erg s $^{-1}$. The data have been corrected for ‘completeness’ (Fabbiano, 2006). This is because instrumental effects and source confusion decrease the efficiency of detection for threshold sources, and unless corrected, the luminosity function will appear too flat at the low-luminosity end.

Figure 13.19 (left) shows a composite XLF made from Chandra observations of 14 E and S0 galaxies containing ≈ 900 detected sources (Kim & Fabbiano, 2004). The best fit is a broken power law with slope -0.8 at low L_x , a break at 5×10^{38} erg s $^{-1}$ and a slope of -1.8 at high L_x . The right side of the figure shows a composite XLF made from Chandra observations of eight starburst galaxies containing ≈ 1000 detected sources (Grimm *et al.*, 2003). This has been fit with a single power law with slope -0.6 and a cut-off at 2×10^{40} ergs s $^{-1}$. Observations of the Small Magellanic Cloud and the Milky Way have been used to determine the function at low L_x . Data from different galaxies have been normalised with a star formation rate derived from UV, H α , FIR and radio fluxes.

Luminosity functions from almost all galaxies show the LMXB/HMXB difference illustrated in Fig. 13.19. Sources from regions containing old

stars are LMXBs, and their maximum L_X is generally a few $\times 10^{38}$ erg s $^{-1}$. This is about the Eddington limit for a $3 M_\odot$ neutron star (the maximum mass possible for a neutron star; see Chapter 12). Sources with higher luminosities are thought to be black holes. Sources from star-forming regions are a mixture of LMXBs and HMXBs, and the XLF often extends beyond 10^{39} erg s $^{-1}$. The bright HMXBs are thought to be black-hole ULXs. The number of LMXBs in a galaxy, in most cases, is proportional to the optical or IR flux – the number of stars. The number of HMXBs is proportional to the star formation rate of the galaxy.

13.10.3 Supersoft and ultra-luminous sources

In addition to finding hundreds of XRBs, observations of other galaxies have uncovered two new types of X-ray source; the supersoft source (SSS) and the ultra-luminous source (ULX). The SSSs are bright, radiate no energy above 1 keV and are typically characterised by a low temperature, $kT \approx 70$ eV. Because an interstellar column density of a few $\times 10^{21}$ cm $^{-2}$ will absorb all energy from the source, observations of small or face-on galaxies are more favourable for SSS detection. This is why the first SSSs, found by Einstein, were CAL 83 and CAL 87 in the LMC. Fifty-seven SSSs were catalogued using ROSAT and Einstein data, all but one in the Local Group (Greiner, 2000), with X-ray luminosities ranging from 10^{36} to 10^{38} erg s $^{-1}$. The SSS are identified with accreting white dwarf stars – some associated with novae. For more details, see Chapter 10.

An ultra-luminous X-ray source (ULX) is point-like, has a luminosity in excess of 10^{39} ergs s $^{-1}$ and is *not* the nucleus of a galaxy. These are found on average of one per galaxy, although some galaxies have several. There is no strong correlation in number or luminosity with properties of the galaxies, except for starburst galaxies. They are associated with young star-forming regions and show a preference for young O associations. There are apparently none in the Milky Way. The first were found by Einstein in the galaxies M81, M82 and M101 (Long & van Speybroeck, 1983).

The Eddington limit for a spherical accreting object of mass M is $1.3 \times 10^{38} M/M_\odot$ ergs s $^{-1}$. If the accreting object is a $1.4 M_\odot$ neutron star,

the limit is $\approx 2 \times 10^{38}$ ergs s $^{-1}$. One possibility, then, for the nature of a ULX is that the accreting object is a black hole with mass $\geq 7 M_\odot$. Other possibilities are non-uniform emission, such as a jet which radiates strongly over a narrow cone, or super-Eddington accretion, in which e.g. an object accretes at the equator and radiates at the poles. All these situations are possible, but it is the high-mass primary that is of most interest. The most luminous ULX is source X-1 in the galaxy M82, which has $L_X = 1 \times 10^{41}$ ergs s $^{-1}$, implying a mass of $70 M_\odot$ if it is Eddington limited. We note that ‘stellar black holes’ in Local Group binary systems have been accurately determined dynamically to be in the range $3\text{--}15 M_\odot$. A bright ULX such as M82 X-1 might therefore signal the presence of an ‘intermediate-mass’ black hole (IMBH) – an object with mass too large to be made by normal stellar evolution, yet smaller than the black holes at the centres of galaxies. This is exciting, but the bright ULXs cannot be taken as definite evidence for IMBHs until alternate explanations have been ruled out and, ideally, dynamically measured masses have been obtained.

We also note that a 10^{41} ergs s $^{-1}$ source can be detected under the conditions of Table 13.2 at a distance of 1000 Mpc. However, although ULXs are easy to detect, they are not easy to identify. A background AGN will have observed properties which cannot be distinguished from that of a ULX without an optical counterpart or a detailed X-ray spectrum, both very difficult to obtain for very distant sources.

References

Observations

- Antoniou, V., et al. (2009) The Chandra survey of the SMC ‘Bar’ II. Optical counterparts of X-ray sources, *ApJ* **697**, 1995.
- Baganoff, F. K., et al. (2001) Rapid X-ray flaring from the direction of the supermassive black hole at the Galactic Center, *Nature* **413**, 45.
- Baganoff, F., et al. (2003) Chandra X-ray spectroscopic imaging of Sgr A* and the central parsec of the Galaxy, *ApJ* **591**, 891.
- Bowyer, S., et al. (1974) Observation of X-rays from M31, *ApJ* **190**, 285.

- Fabbiano, G., A. Zezas & S. Murray (2001) Chandra observation of 'The Antennae' galaxies (NGC 4038/39), *ApJ* **554**, 1035.
- Gaetz, T. J., et al. (2007) Chandra ACIS survey of M33 (ChASEM33): X-ray imaging spectroscopy of M33SNR21, the brightest X-ray supernova remnant in M33, *ApJ* **633**, 234.
- Garcia, M., et al. (2005) A possible detection of M31* with Chandra, *ApJ* **632**, 104.
- Greiner, J. (2000) Catalog of supersoft X-ray sources, *New Astron* **5**, 137.
- Grimm, H.-J., M. Gilfanov & R. Sunyaev (2003) High mass X-ray binaries as a star formation rate indicator in distant galaxies, *MNRAS* **339**, 793.
- Kahabka, P., et al. (1999) A ROSAT PSPC X-ray survey of the Small Magellanic Cloud, *A&AS* **136**, 81.
- Kim, D.-W. & G. Fabbiano (2004) X-ray luminosity function and total luminosity of LMXB in early type galaxies, *ApJ* **611**, 846.
- Kim, D.-W., et al. (2006) Probing the low-luminosity X-ray luminosity function in normal elliptical galaxies, *ApJ* **652**, 1090.
- Kong, A., et al. (2002) X-ray point sources in the central region of M31 as seen by Chandra, *ApJ* **577**, 738.
- Leong, C., et al. (1971) X-ray emission from the Magellanic Clouds observed by Uhuru, *ApJ* **170**, L67.
- Long, K. & L. van Speybroeck (1983) X-ray emission from normal galaxies, in *Accretion-Driven Stellar X-ray Sources*, 117, edited by W. Lewin & E. van den Heuvel. Cambridge: Cambridge University Press.
- Mark, H., et al. (1969) Detection of X-rays from the Large Magellanic Cloud, *ApJ* **155**, L143.
- Muno, M. P., et al. (2003) A deep Chandra catalog of X-ray point sources toward the galactic center, *ApJ* **589**, 225.
- MPE/ROSAT (1991) <http://www.mpe.mpg.de/xray/wave/rosat/gallery/calendar/1991/sep.php>.
- MPE/ROSAT (1992) www.mpe.mpg.de/xray/wave/rosat/gallery/calendar/1992/oct.php.
- NASA/CXC (2002) <http://chandra.harvard.edu/photo/2002/gcenter/index.html>.
- NASA/CXC (2003) <http://chandra.harvard.edu/photo/2003/1154/>.
- NASA/CXC (2006) http://chandra.harvard.edu/photo/2006/m31/m31_xray_opt_pullout.jpg.
- NASA/CXC (2007) <http://hea-www.harvard.edu/ChandraSNR/>.
- NASA/SAO (Fabbiano et al. (2001) <http://chandra.harvard.edu/photo/2001/0094true/>.
- Orosz, J., et al. (2007) A 15.65-solar-mass black hole in an eclipsing binary in the nearby spiral galaxy M33, *Nature* **449**, 872.
- Pietsch, W., et al. (2004) XMM-Newton survey of M33, *A&A* **426**, 11.
- Pietsch, W., M. Freyberg & F. Haberl (2005) An XMM-Newton survey of M 31, *A&A* **434**, 483.
- Prestwich, A., et al. (2003) Classifying X-ray sources in external galaxies from X-ray colors, *ApJ* **595**, 719.
- Price, R., et al. (1971) X-rays from the Magellanic Clouds, *ApJ* **168**, L7.
- Randall, S., C. Sarazin & J. Irwin (2004) Chandra observation of diffuse gas and LMXBs in the elliptical galaxy NGC 4649 (M60), *ApJ* **600**, 729.
- Read, A. M. & I. R. Stevens (2002) XMM-Newton reflection grating spectrometer observations of the prototypical starburst galaxy M82, *MNRAS* **335**, L36.
- Skinner, G., et al. (1987) Hard X-ray images of the galactic center, *Nature* **330**, 544.
- Soria, R. & K. Wu (2003) Properties of discrete X-ray sources in the starburst spiral galaxy M83, *A&A* **410**, 53.
- Stevens, I., A. Read & J. Bravo-Guerro (2004) XMM-Newton EPIC observations of the prototypical starburst galaxy m82, *Mem Soc Astron Ital* **75**, 513.
- Supper, R., et al. (1997) ROSAT PSPC survey of M 31, *A&A* **317**, 328.
- van der Hayden, K., J. Bleeker & J. Kaastra (2004) Synoptic study of the SMC SNRs using XMM-Newton, *A&A* **421**, 1031.
- van Speybroeck, L., et al. (1979) Observations of X-ray sources in M31, *ApJ* **234**, L45.
- Vargas, M., et al. (1996) SIGMA survey of the Galactic Center region and discovery of a hard X-ray transient, *A&AS* **120**, 291.
- Wang, Q. D., E. V. Gotthelf & C. C. Lang (2002) A faint discrete source origin for the highly ionised iron emission from the Galactic Center region, *Nature* **415**, 145.
- Williams, R., et al. (2005) Supernova remnants in the Magellanic Clouds. V. The complex interior structure of the N206 supernova remnant, *ApJ* **628**, 704.

Catalogues and reviews

- Fabbiano, G. (2006) Populations of X-ray sources in galaxies, *Ann Rev Ast Ap* **44**, 323.
- Haberl, F. & W. Pietsch (1999) A ROSAT PSPC catalogue of X-ray sources in the LMC region, *A&AS* **139**, 277.
- Haberl, F. & W. Pietsch (2001) The X-ray view of M33 after ROSAT, *A&A* **373**, 438.
- Haberl, F., et al. (2000) A ROSAT PSPC catalogue of X-ray sources in the SMC region, *A&AS* **142**, 41.
- STScI/DSS (2007) <http://archive.stsci.edu/cgi-bin/dss-form>.

Chapter 14

Active galactic nuclei

14.1 Introduction

Normal galaxies like our own, when viewed from great distances, appear to be peaceful and unchanging aggregations of stars, whose well-being is only slightly disturbed by the occasional supernova explosion. However, violent processes far more powerful than supernovae have been known since early in this century. The optical jet emanating from the giant elliptical galaxy M87 (the dominant galaxy in the relatively nearby Virgo cluster of galaxies) was found in 1917, but its significance was not understood for many years. After the Second World War, the founding of radio astronomy led to the discovery of luminous extra-galactic radio sources such as Cygnus A. Also, short-exposure optical photographs showed that some apparently normal spiral galaxies actually had very bright, almost starlike nuclei, the prime example of which is NGC 4151 (Fig. 14.1), hence the term *active galactic nuclei*, or AGN.

Such galaxies are referred to as Seyfert galaxies, after their discoverer, Carl Seyfert. But even these exotic objects paled in comparison with the enormous energy output at all wavelengths of quasi-stellar objects (better known as quasars, or QSOs), discovered originally through their radio emission in the early 1960s and so-called because of their ‘stellar’ appearance. However, the discovery in 1963 of their very high red-shifts (see Box 14.1) implied that QSOs were immensely distant, and hence they were the most luminous objects in the Universe (see Fig. 14.2). QSOs were also found to be variable, some on timescales of a

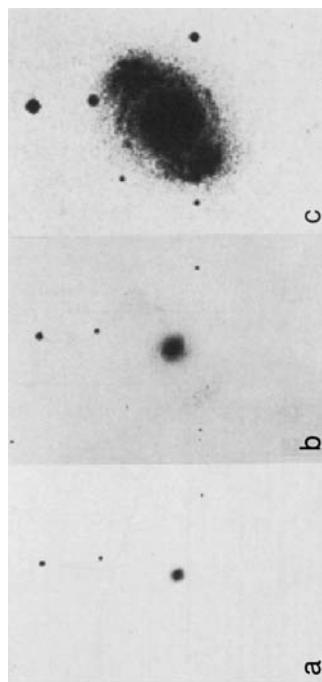


Fig 14.1 Short, medium and long exposures of NGC4151, showing the bright starlike nucleus and the surrounding spiral galaxy.

few hours, which meant that the region responsible for this enormous emission was incredibly tiny – essentially, 1000 times the output of our entire Milky Way galaxy was coming from a volume not much bigger than our Solar System! One of the great theoretical astrophysics problems of the 1960s was to explain the fundamental nature of QSOs.

Box 14.1 H_0 and the expansion of the Universe

The distances of most of the extragalactic objects referred to in this book are sufficiently small that we can estimate their distance, and hence their intrinsic luminosity, by the methods of classical astronomy. In addition, Hubble had shown in the 1920s that there was a linear relationship between the distance of a galaxy and its velocity of recession from us. This is Hubble's law and led to the concept of an expanding Universe. The constant of proportionality in this relationship has been determined by HST to be $72 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and is known as Hubble's constant, or H_0 .

The observed red-shift of an object is defined quite straightforwardly as $z = (\lambda_{\text{obs}} - \lambda)/\lambda$, the ratio of the wavelength shift of a particular feature (e.g. emission lines of hydrogen) to the rest wavelength of that feature. However, to relate z to the line-of-sight velocity of the object is not straightforward. At low velocities ($z \leq 0.1$), it is $z \approx v/c$, but at larger values of z , this breaks down (especially now that z values exceeding 8 have been observed). The full relativistic expression is

$$\frac{v}{c} = \frac{(1+z)^2 - 1}{(1+z)^2 + 1}, \quad (14.1)$$

from which the following values can be obtained:

z	v/c
0	0
0.05	0.05
0.10	0.10
0.50	0.38
1.00	0.60
2.00	0.80
5.00	0.95
10.00	0.98

The physical conditions implied by the creation of such high luminosities within a small volume led to a long and highly controversial debate. Some doubted that quasar red shifts were being correctly interpreted. Perhaps the red shift was due to physical conditions near the quasar rather

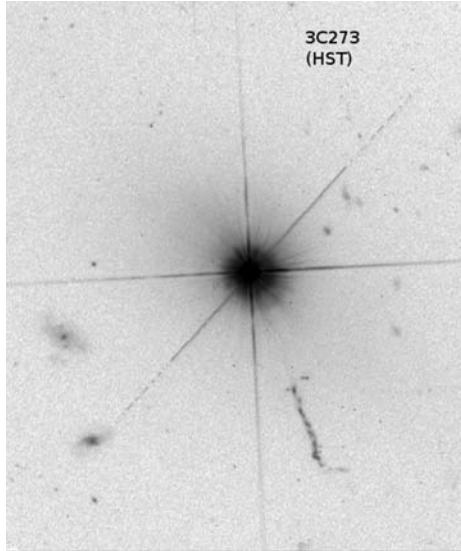


Fig 14.2 HST optical image of the prototype QSO, 3C273, with its remarkable jet extending about 10 arcsec to the southwest. 3C273 is the closest (750 Mpc) and brightest (13th magnitude) QSO, and the extreme brightness of the nucleus is demonstrated by comparing it with 'normal' galaxies at the same distance, which are easily visible in this image (e.g. bottom left region) and are all part of the same cluster of galaxies. Image courtesy NASA/STScI.

than to cosmological distances. However, the nearest quasars were subsequently found (when examined with large ground-based telescopes and then more impressively with HST) to be embedded in extended objects with features akin to spiral galaxies. The red-shifts of these extended regions (and of surrounding galaxies, such as in Fig. 14.2) were found to be identical to those of the quasar, thereby providing overwhelming support for the cosmological interpretation of quasar red shifts, which is now almost universally accepted. Thus quasars can be considered extreme examples of Seyfert galaxies, and a selection of HST images of QSOs is shown in Fig. 14.3. Indeed, at the other end of the scale, even local, normal galaxies like our own are now recognised to exhibit low levels of nuclear activity, indicating that such properties are ubiquitous. It is only their level of activity that changes. The determination of these physical properties is the subject matter of this chapter.

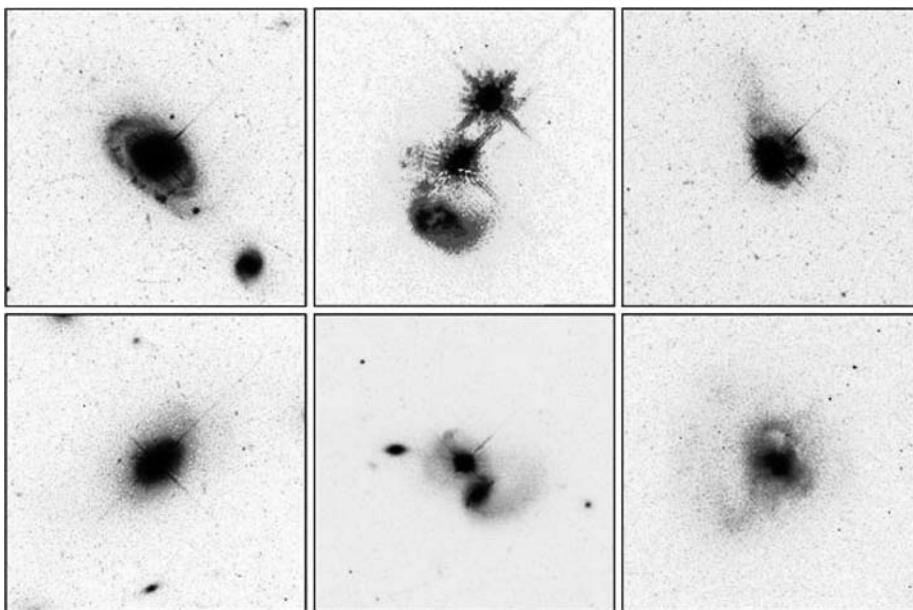


Fig 14.3 Images of QSOs that exploit HST's superb spatial resolution to reveal the details of the surrounding host galaxies (courtesy NASA).

14.2 The range of activity in 'Active' galaxies

The similarity of the optical spectrum of a Seyfert galaxy nucleus to that of a quasar implies that both are the product of the same mechanism. Indeed, if a Seyfert galaxy were to be moved sufficiently far away that the host spiral galaxy is no longer detectable, we would classify it as a quasar. As alluded to in the previous section, the discovery of extended 'fuzz' (the host galaxy) around quasars reinforces this view (Fig. 14.3), and it is not surprising that the highest-luminosity (most distant) Seyferts overlap with the lowest-luminosity (closest) quasars. This range of activity is well demonstrated by the optical spectra shown in Fig. 14.4. Table 14.1 summarises the X-ray, optical and radio properties of the complete range of active galactic nuclei and gives them with respect to normal galaxies. Only the nearest and brightest AGN were detected with the early, non-imaging X-ray surveys, but this changed substantially post-Einstein, and the Chandra/XMM surveys of selected regions are revealing huge numbers of AGN.

14.2.1 Deep X-ray surveys with Chandra and XMM

Two regions of the sky (one in the north, one in the south, designated CDF-N and CDF-S, respectively) have received very deep exposures of ≈ 2 Ms with Chandra. These were chosen to be analogous to the famous Hubble Deep Fields in the optical, both of which are included within the Chandra fields. The CDF-N image and the corresponding Hubble identifications are shown in Colour Plates 54 and 55.

These regions were chosen because of (1) their absence of bright (nearby) stars and galaxies and (2) their high galactic latitudes (and hence low extinction), which would otherwise have obscured distant objects. The combination of Chandra's exquisite angular resolution and sensitive X-ray detectors reveals an AGN spatial density of ~ 7200 sources/deg². These images are clearly dominated by point sources and contrast dramatically with the Hubble images, which contain mostly normal, albeit young and still forming, galaxies. This indicates that the X-ray images show the 'black hole' light emanating from the AGN, and at the highest z (where it would be impossible to resolve the AGN light from that of the host) the X-ray

Table 14.1 Properties of AGN.

Type	L_X (ergs s $^{-1}$)	Host galaxy	Radio emission ^a	Example
BL Lac objects	10^{45-49}	Elliptical	RL	OJ287
QSOs	10^{45-47}	Interacting spirals	RL, RQ	3C273
Seyfert galaxy	10^{43-45}	Spiral	RQ	NGC4151
Narrow em.l. galaxy	10^{40-43}	–	RQ	Mkn176
Normal galaxy	10^{37-39}	–	RQ	M31

^a RL = radio loud; RQ = radio quiet.

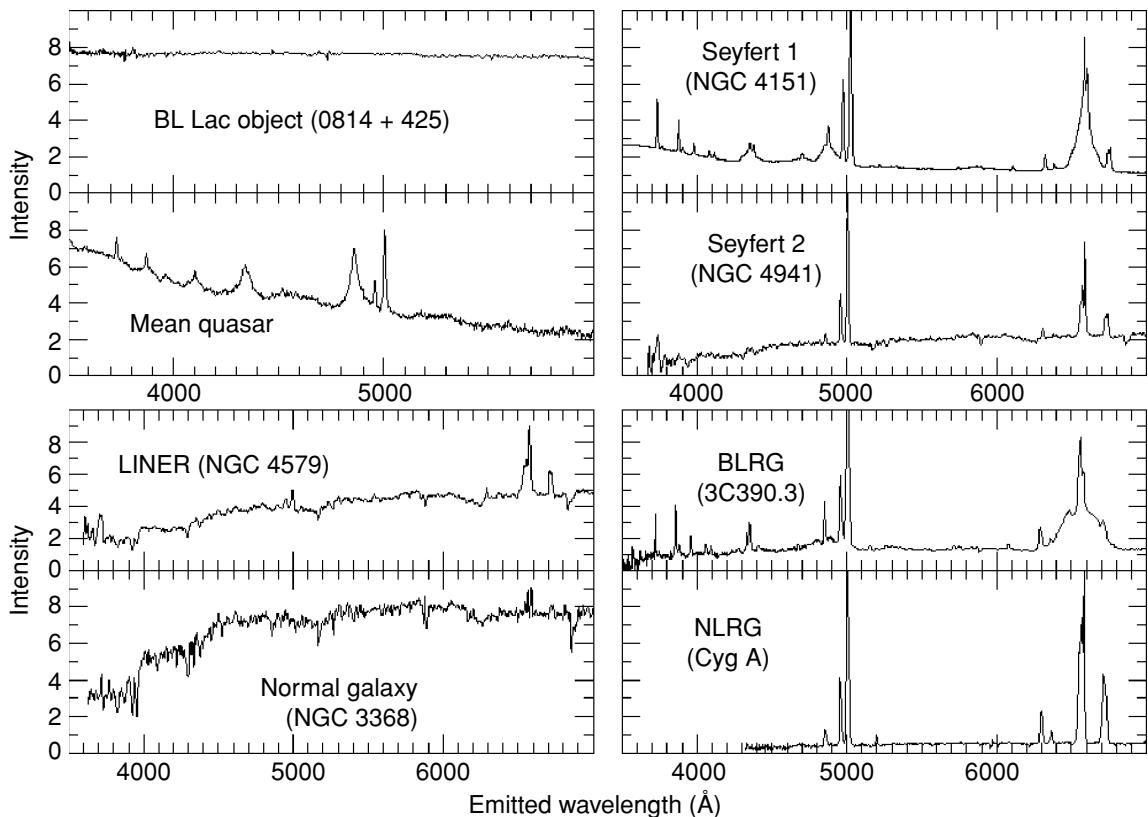


Fig 14.4 Typical optical spectra of AGN, with examples here covering the full range from the (essentially featureless) BL Lac objects (top left) through the radio galaxies and quasars. Note the presence of both broad and narrow components in quasars, Seyfert 1 galaxies and broad-line radio galaxies (BLRG), which arise in completely different regions surrounding the core of the galaxy. However, the broad component is completely missing in Seyfert 2 galaxies, low-ionisation nuclear emission region (LINER) galaxies and narrow-line radio galaxies (NLRG), which only contain narrow permitted and forbidden lines. A normal galaxy spectrum (bottom left) is shown for comparison, and the continuum and weak absorption features of normal stars can be seen from the underlying host galaxy in Seyfert 2 and LINER galaxies.

observations maximise the contrast against background emission from the host. Furthermore, subsequent optical/radio follow-up observations show that >75 per cent of all the sources in the CDFs are AGN, with red shifts in the range 0.1–5.

Remarkably, the deep X-ray surveys show that the properties of these AGN are roughly constant from $z \sim 0-6$ and that there is no evidence that any (substantially) exceed the Eddington limit (see Box 14.3).

14.3 The unified model of AGN: What is the central engine?

When QSOs were discovered in the 1960s, it was recognised that an explanation of their properties required an extraordinary central engine that was capable of generating enormous amounts of power in only a small volume of space. At that time, the most powerful astronomical objects known were supernovae, and at least 10 000 of them (all at their peak brightness) would be required to account for the output of typical QSOs. The number of stars required to produce and maintain this number of constantly occurring supernovae quickly ruled out this explanation.

Large numbers ($>10^{10}$) of normal stars could also be excluded from consideration because of the short-term variability observed in QSOs and the small volume in which they would have to exist. Furthermore, normal stars are not very efficient at converting matter into energy via normal fusion processes (<1%). However, even though not a single such object had been observed at that time, several theorists (Salpeter, Novikov, Zel'dovich) recognised that accretion onto a supermassive black hole (in the range 10^6 – $10^{10} M_\odot$) would be capable of generating QSO-level luminosities because of the much greater energy-generating efficiency of the accretion process (see Chapter 11). This was taken further by Lynden-Bell (1969), who established the basic structure of the accretion disc that would form in such a process. At the time these were considered radical, even extreme, concepts, but the last 30 years have seen an enormous amount of observational evidence that has led to the wide acceptance of this model.

It is instructive at this point to consider the similarity of certain radio galaxy jets to the much smaller scale versions seen in SS433 and other microquasars within the Milky Way (see Chapter 12). The similarity of their properties is striking and demonstrates that accretion power must be the key. 3C129 is an excellent example of this, in which a precessing beam has been ‘blown’ backwards by passage through the intracluster medium (Fig. 14.5). It is simply a larger-scale version of the model that we envisage for SS433. Just

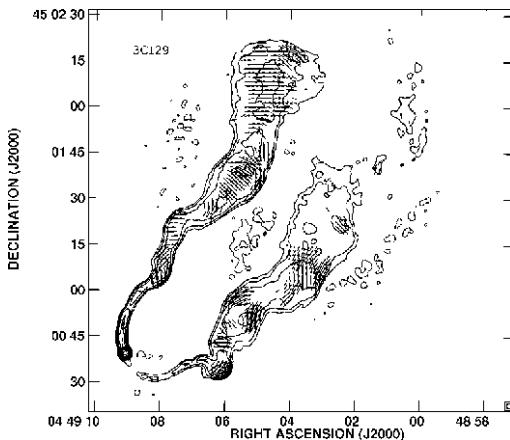


Fig 14.5 The radio galaxy 3C129 exhibits radio features that are the result of a precessing jet from the nucleus which is ‘swept back’ by ram pressure created by the radio galaxy as it travels through the intracluster medium. In fact, this observation provides direct evidence for the existence of the medium, which is, despite its low density, very hot and produces thermal X-rays which we see as diffuse emission from the cluster as a whole (as discussed in Chapter 15) (Taylor et al., 2001).

compare this radio image of 3C129 with Fig. 11.34, showing SS433’s precessing radio jets.

While microquasars have only been identified relatively recently, it is becoming clear that there are many analogies between them and full-blown QSOs. We know that the basic energy source in SS433 is accretion via an interacting binary system (Chapter 16) and that SS433 is suffering from exceptionally high mass-transfer rates. For reasons still not fully understood, a significant fraction of the material is not accreted but is instead ejected as a jet along the spin axis of the compact object at a velocity of 0.27c. These jets precess with a 160-day period about the (presumed) spin axis of the disc. While SS433 is a stellar-sized object producing ‘only’ about 10^{36} ergs s⁻¹ at X-ray wavelengths, its total power in the jets is at least 1000 times greater. This basic model can be directly applied to QSOs by simply scaling it up by the mass of the compact object (see Fig. 14.6).

To extract from the system a factor of a million times more energy requires the mass of the compact object to be greater by exactly the same factor (see Box 11.2). Such an object must be

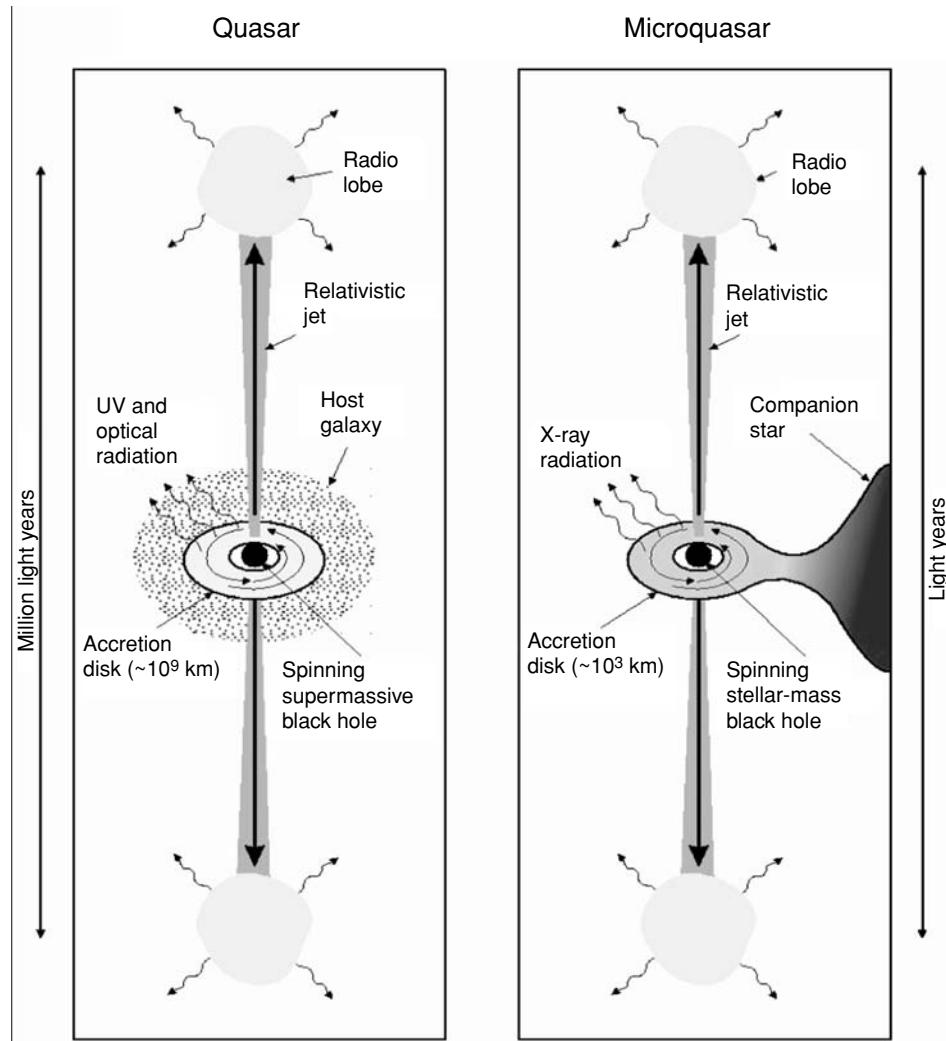


Fig 14.6 The basic physical processes shown here in (left) QSOs and (right) galactic X-ray binaries are believed to be identical. They differ only in terms of scale, which is driven by the mass of the compact object (figure by Mirabel, 2007).

a black hole. However, the fuel for the accretion mechanism must also be found. We can no longer consider a single mass-losing star, as in the X-ray binaries, but must search for sources of gas on much larger scales. Powerful quasars require accretion rates of several solar masses per year to explain their prodigious energy output. Although the higher density of normal interstellar material (from supernovae, novae and stellar winds) near the galaxy's centre may account for some of the fuel, it is likely that the higher stellar density combined with the black hole's intense gravitational

field could tidally disrupt entire stars and generate a massive disc around the hole. Such a system is shown schematically in Fig. 14.6.

Out of this has emerged the ‘unified’ model of AGN, in which many of their observed properties can be explained with a ‘standard’ central engine which is simply viewed at different orientation angles to our line of sight. The different types of AGN are the result of different viewing directions. This is shown schematically in Colour Plate 39, where the central disc (or torus), which supplies the material for accretion, acts to block

our line of sight in the plane of the disc. Since the most rapidly orbiting clouds of gas (which are responsible for the broad optical emission lines) are inside the torus, they can only be viewed if the observer is at a high inclination. Such an AGN would be catalogued as a Seyfert 1 galaxy. However, at lower inclinations, the broad-line region (BLR) is hidden by the torus, leaving only the much slower moving, more extended narrow-line region (NLR) visible, which we would describe as a Seyfert 2 galaxy.

This unified model for Seyfert galaxy types received dramatic support in 1985 when Antonucci and Miller observed Seyfert 2 galaxies to see how the degree of polarisation of their light varied with wavelength. They discovered that the polarisation spectra were essentially those of Seyfert 1 galaxies, i.e. they contained *broad* lines which were not visible in the direct spectra. This is exactly what you would expect if Seyfert 2 galaxies were the same as Seyfert 1 galaxies, but simply viewed in the plane of the torus. The emission from the BLR would only be seen in the scattered (and hence polarised) light from material (dust or electrons) well outside the torus. Furthermore, Seyfert 2 galaxies have lower X-ray luminosities and higher absorbing column densities (by a factor of 10–100) than Seyfert 1 galaxies, exactly as would be expected in the unified model.

14.4 Evidence for the existence of supermassive black holes

These arguments, which have led to almost universal acceptance that the central engine in AGN is an accreting supermassive black hole, are compelling but for many years were based largely on circumstantial evidence. Nobody had actually measured the mass of the central object in an AGN until the 1980s, and the most direct evidence has appeared only in the last 15 years or so. All mass measurements are based on locating material that is as close as possible to the central object and (1) whose motion around the object can be directly observed, thereby providing for dynamical measurements of the BH masses or (2) whose emission spectra

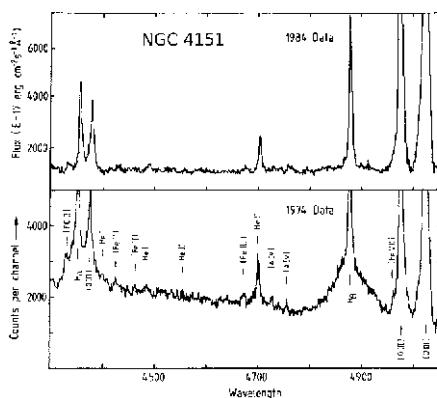


Fig 14.7 Optical spectra of NGC4151 obtained with the INT in 1974 and 1984. The activity of the nucleus has changed dramatically in 10 years from that of a classic Seyfert 1 to the much milder Seyfert 2 (in which the broad-line emission corresponding to the highest-velocity material is greatly reduced) (spectra by Mike Penston).

display relativistic distortions due to the presence of the compact object.

14.4.1 Reverberation: NGC4151

As is often the case in astronomy, exotic ideas, such as supermassive black holes ejecting relativistic jets of material over vast reaches of space, are often best tested by studying related but less extreme objects on our own ‘doorstep’. The prototype (and brightest) Seyfert galaxy, NGC4151 (Fig. 14.1), is certainly an interesting and striking object, and its brightness makes it possible to obtain high-quality data in reasonable observing times. The late Mike Penston (RGO) and his group of *old LAGs* (Lovers of Active Galaxies) made one of the earliest direct estimates of an AGN central mass.

They observed NGC4151 from the ground and in space for almost 10 years. One of the startling results of their monitoring is shown in Fig. 14.7, which contains optical spectra of the galaxy’s nucleus taken with the 2.5-m Isaac Newton Telescope in 1974 (when it was based in Sussex at Herstmonceux Castle) and 1984 (from its new La Palma site). The enormously broad emission line component that signified its classification as a Seyfert 1 galaxy (the most active variety) had vanished. If NGC4151 had been observed, then, for the first time, it would have been classified as the much

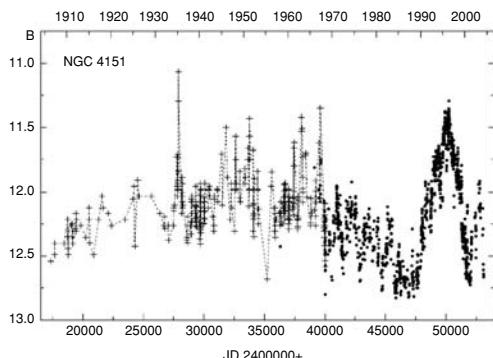


Fig 14.8 A \sim 100-year visual (B-band) light curve of the ‘stellar’ component at the core of the Seyfert galaxy NGC 4151. Its short- and long-term variations are clearly evident, and these are detectable at all wavelengths, but with time lags between them.

less active Seyfert 2. Clearly these changes had to be taking place close to the core of the galaxy.

Observations in the far UV with IUE gave more detail on the nature of the ‘clouds’ of rapidly moving gas in the core region. The variability of the ‘stellar’ core of NGC 4151 is well known (see Fig. 14.8), and in 1979, Penston and his team were fortunate to observe a brightening of the core (the unresolved stellar component) but noted that the clouds did not increase in brightness until approximately 13 days later. There is clearly a time lag between the variation of the continuum (produced in the core) and the emission lines (that originate in the surrounding clouds).

This reverberation in the central region of NGC 4151 presented a remarkable opportunity to determine some of the basic physical parameters of its nucleus if certain simplifying assumptions were made. The detailed calculation is laid out in Box 14.2, but essentially, we are assuming that the ‘clouds’ are orbiting the central core in exactly the same way that the planets orbit the Sun. The observed time delay directly yields the distance of the clouds from the core, and hence the mass of the core can be estimated. It is about $10^9 M_\odot$, a truly supermassive object, which must be a black hole. NGC4151’s activity is noticeable to us because of its proximity, and there is no evidence at present for radio jets, as in the more powerful quasars.

Box 14.2 Weighing the black hole in NGC4151

The basis of this calculation is shown schematically in Fig. 14.9. A supermassive central object (assumed to be a black hole) of mass M is the ‘central engine’ that generates the blue continuum radiation. This radiation excites the material in the surrounding clouds, thereby producing the strong emission lines. The clouds are at a distance r from the central object, which they orbit with a velocity v .

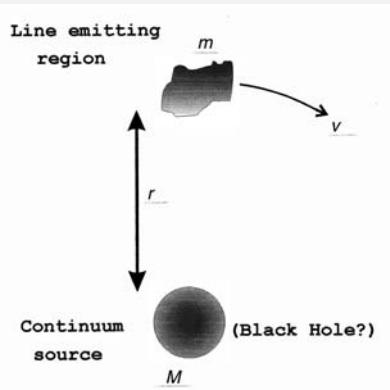


Fig 14.9 How to weigh the presumed black hole at the centre of an active galaxy.

When the UV continuum radiation from NGC4151 was observed to brighten, there was a 13-day delay before the emission lines brightened too (the reverberation timescale). This means that r is 13 light days, or 3.4×10^{16} cm. The observed widths of the emission lines (e.g. CIV) give $v = 14000 \text{ km s}^{-1}$, or $1.4 \times 10^9 \text{ cm s}^{-1}$. We now assume that this velocity is the circular velocity required to maintain the cloud in orbit about the black hole. In other words,

$$\frac{mv^2}{r} = \frac{GMm}{r^2}, \quad (14.2)$$

which simplifies to give $M = v^2r/G$, and inserting the preceding numerical values yields $M = 10^{42} \text{ gm}$, or 500 million M_\odot !

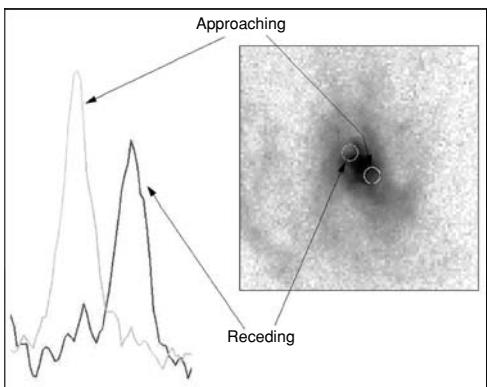


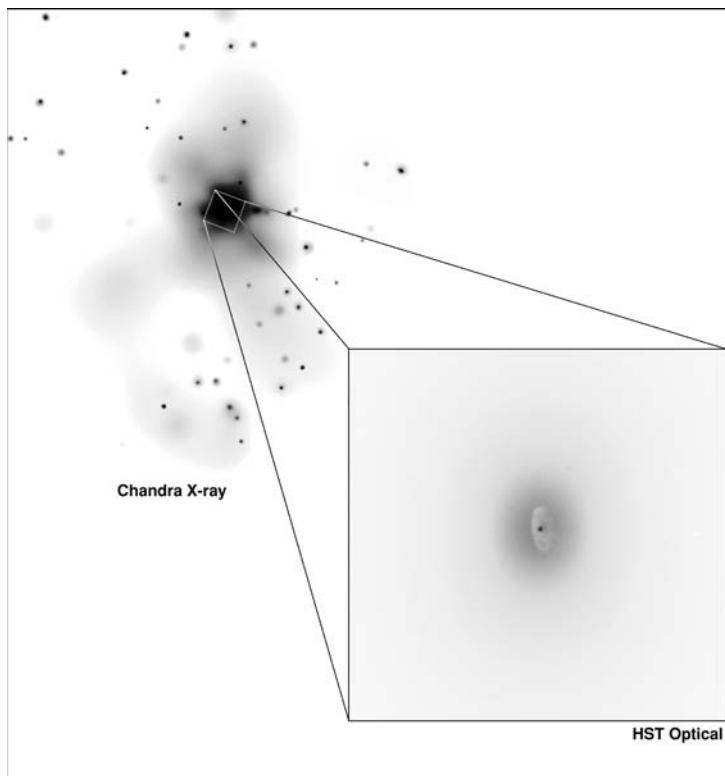
Fig 14.10 (right) HST narrowband H α image of the core region of M87 which reveals the presence of the extended accretion disc. (left) HST FOS spectra of the circled regions show blue- and red-shifted emission profiles of H α from which the central object mass can be estimated, assuming that the gas in the disc is in Keplerian motion about the centre (image courtesy of NASA).

14.4.2 Resolving the disc with HST

The unified model for AGN requires the existence of a huge accretion disc surrounding the SMBH, and attempts to directly image such a disc were finally successful with HST. In 1994, Holland Ford and co-workers used HST to both image and obtain spectra (Fig. 14.10) of the core region of M87, the giant elliptical galaxy in the Virgo cluster of galaxies. The spectra indicated gas velocities of $\sim 460 \text{ km s}^{-1}$ at an angular separation of just 0.25 arcsec from the centre (equivalent to 20 pc at the distance of M87). As for the reverberation calculation for NGC 4151 (Box 14.2) it is assumed that the gas in the disc is in Keplerian motion around the central object, so its mass is $M = v^2 r / G$. Putting in the observed values of v and r gives $M_{BH} \sim 3 \times 10^9 M_\odot$, a truly supermassive black hole and the nearest one to us.

An HST image of the radio galaxy NGC 4261 is shown in Fig. 14.11. This figure includes the Chandra X-ray image which shows the bright QSO core surrounded by a population of ‘classical’ X-ray sources, most of which are stellar-mass X-ray

Fig 14.11 Chandra and HST images of the radio galaxy NGC 4261. The disc of the luminous central X-ray source is resolved by the HST image (composite image courtesy of NASA).



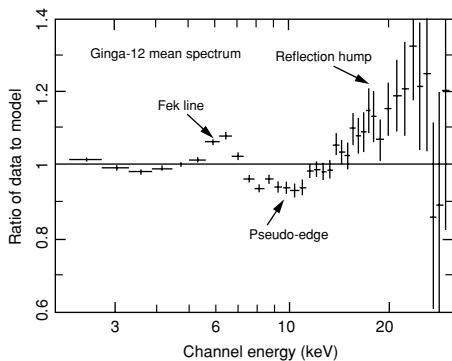


Fig 14.12 Ginga-12, the average of the X-ray spectra of 12 AGN observed with the Ginga satellite. The spectrum is plotted here as a ratio against a simple power law to demonstrate the additional features. These include the iron line between 6 and 7 keV, an apparent edge just above this and then the Compton reflection component (figure adapted from Pounds *et al.*, 1990).

binaries. The HST-resolved disc has a remarkable similarity to the unified model schematic, and similar HST spectra of the extended disc region reveal a central compact object mass approximately two-thirds that of M87. Clearly SMBHs are a feature of galaxies that exhibit radio jets and lobes.

14.4.3 Iron in X-ray spectra: The effects of strong gravity

14.4.3.1 Early X-ray spectroscopy of AGN
 It is hard to find a better example of the spectacular improvements in X-ray astronomy technology than through our knowledge of the X-ray spectra of AGN. In the early days of X-ray astronomy (the Uhuru mission), non-imaging X-ray proportional counter detectors could barely detect AGN, let alone measure their spectra. The sensitivity improved with EXOSAT and Ginga as it became possible to build larger detectors, but the spectral resolution remained the same. Nevertheless, it had been recognised by the early 1990s that QSO X-ray spectra (from low-luminosity Seyfert 2 galaxies to high-luminosity QSOs) could all be well represented by a hard X-ray power law with a slope of -0.7 . This similarity led Ken Pounds and collaborators to compile a ‘mean’ X-ray spectrum of 12 bright Seyfert 1 galaxies with Ginga, and this is shown in Fig. 14.12. The spectra were

of good-enough quality to detect the presence of iron emission between 6 and 7 keV, but surprisingly, the typical energy of this line was 6.4 keV, which is close to the energy expected for the $\text{K}\alpha$ line of neutral iron. This indicates that the emission cannot be thermal in origin as the very hard X-ray spectra imply very high temperatures, which would lead to iron emission at energies of 6.7 keV or higher. Our understanding of this iron line did not improve until the first CCD-based X-ray spectrometer results from the Japanese ASCA satellite in the early 1990s.

14.4.3.2 Getting closer to the event horizon

Impressive though the HST images and spectra are of M87 and NGC 4261, the observations are essentially of regions far removed from the putative black hole. What could be done to probe the regions much closer to the compact object, ideally as close as possible to the event horizon? Consider the properties of the material in the inner regions of the accretion disc. The effects of viscosity within the disc (likely due to magnetic fields trapped therein) result in material losing angular momentum and gradually moving inwards through the disc. In this process, the gas gets hotter and hotter, and the potential energy is converted into radiation. Detailed calculations show that the bulk of the energy released in this way takes place within the very innermost parts of the disc, only ~ 20 gravitational radii or so from the compact object. The associated temperatures are extremely high, with the peak radiation occurring at X-ray wavelengths. Supporting evidence for the X-ray emission arising in such a compact volume of space comes from the X-ray variations. The Seyfert 1 galaxy, MCG-6-30-15, changes by a factor of 2 in just 100 s or so. Nevertheless, this is still circumstantial evidence for the location of the X-ray-emitting region. Direct evidence has recently been gathered by observation of emission line profiles, a technique made possible by technological improvements in X-ray spectroscopy (particularly in sensitivity). The first significant observation was accomplished when Yasuo Tanaka used the CCD spectrometer on board ASCA to observe the X-ray spectrum of MCG-6-30-15 (Fig. 14.13). There was a strong, broad iron line present, with a remarkable, distorted profile.

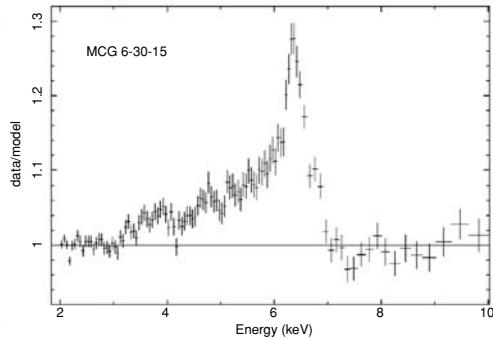
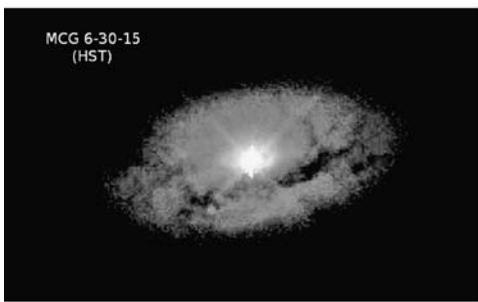


Fig 14.13 (left) The Seyfert I galaxy MCG-6-30-15 as observed with HST, with its bright, active core clearly visible. (right) In X-rays, it displays a distorted Fe K α line profile, as observed with the CCD spectrometer on board the ASCA satellite (image by Matt Malkan, spectrum by Andy Fabian, from Tanaka et al., 1995).

The presence of iron itself is not surprising. Iron is the most abundant of the heavy elements, and it can be a very strong feature in young SNRs (see Chapter 8). However, the line profile was very strange and required a combination of factors to account for it. To understand this, we need to step through the processes that produce the X-rays in these inner regions:

- The extremely high temperatures in the inner disc and the high luminosity of the AGN lead to material being ‘evaporated’ from the surface of the disc and the creation of an extremely hot, tenuous corona.
- This corona radiates hard (many kilo-electron-volts) X-rays in all directions, some towards us, others back towards the disc.
- The inner disc is hot but is sufficiently ‘thick’ that not all the atoms are fully ionised. Iron is able to retain its inner shell electrons at temperatures up to several million degrees, at which point almost all other (abundant) elements are ionised.
- The X-rays that hit the disc are absorbed, some by scattering off electrons but others by the inner (K) shells of iron atoms, putting them in an excited state which decays immediately by the emission of a 6.4-keV X-ray. This process is called *fluorescence*.

Calculations of this process lead to the predicted spectrum shown in Fig. 14.14, where the Fe K α line dominates. However, this profile is sharp and bears little resemblance to that observed.

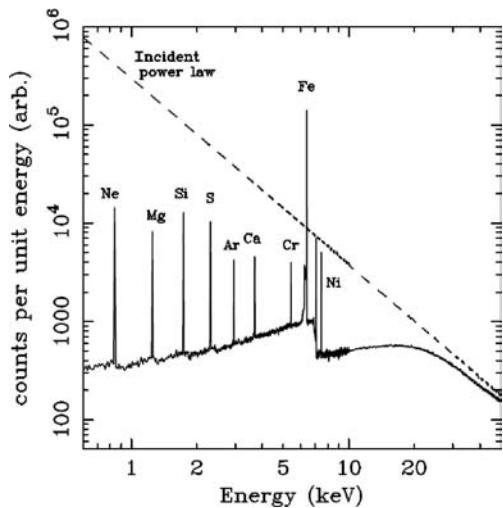
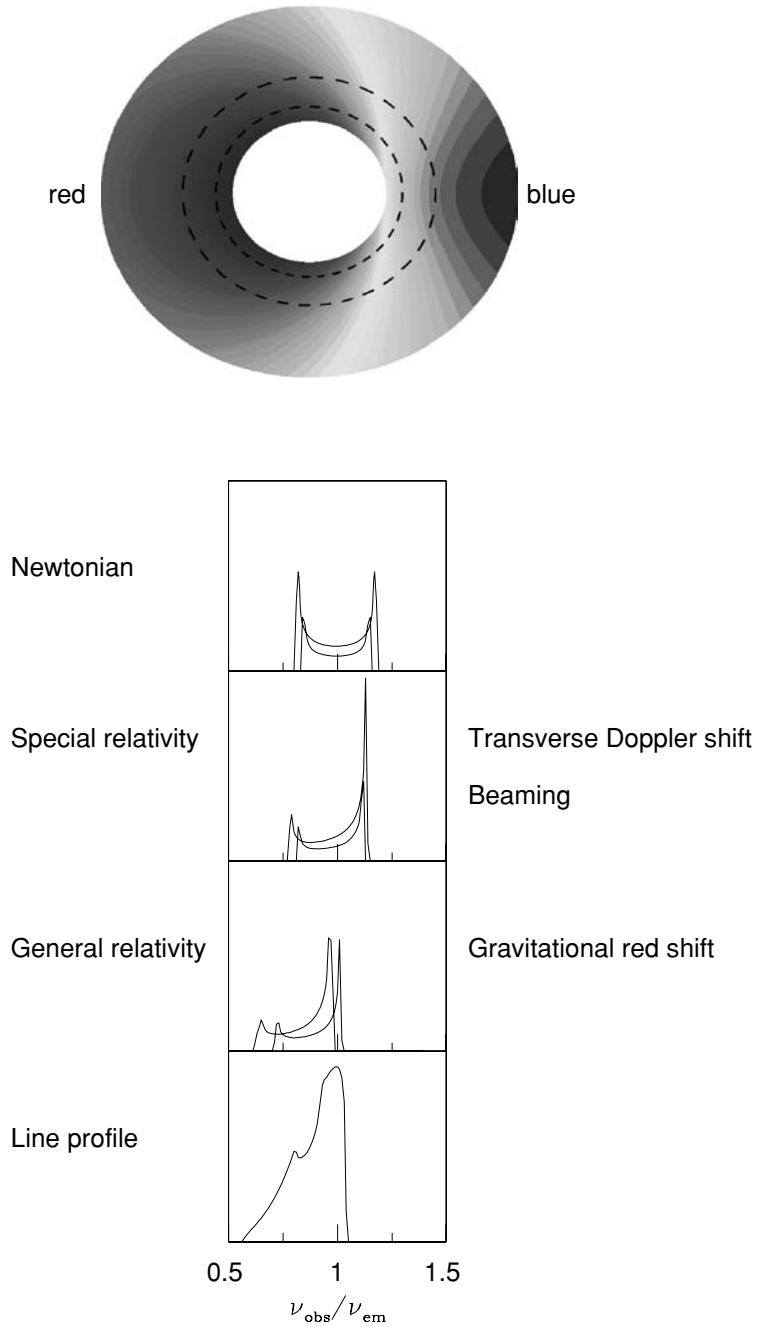


Fig 14.14 Numerical calculation of the effect of irradiating an accretion disc with a hard X-ray spectrum (the incident power law). The resultant ‘reflected’ X-ray spectrum is dominated by iron, the most abundant heavy element, and with the highest fluorescent yield (calculation by Ross & Fabian, 1993).

The material producing the iron profile lies in the very hottest, innermost regions of the disc. Consequently, there are a number of ‘smearing’ effects that must be taken into account when calculating the expected line profile. They are shown schematically in Fig. 14.15 and are as follows:

- *Newtonian mechanics*. The disc rotates faster at smaller radii (exactly as the inner planets of the Solar System orbit much more rapidly than the

Fig 14.15 Idealised view of the inner accretion disc, rotating so that the right-hand side is approaching us (and thereby appears blue shifted). If the disc radiates Fe K α uniformly, then the line will appear split into equal blue and red components due to Newtonian mechanics (top box). But it will be distorted by two effects of special relativity: beaming, whereby matter moving towards us appears brighter than that moving away, and the transverse Doppler shift, which is a bulk shift to the red (second box down). Finally, the disc is sitting in the enormous gravitational well of the central black hole, and so general relativity produces a red-shift of the entire emission (third box down). The sum of these effects, calculated for the whole disc (rather than just the two dotted circles shown) is shown in the bottom box (schematic by Andy Fabian).



outer ones), and we see this as causing the Fe K α line to be split into two peaks, one blue-shifted towards us, the other red-shifted away.

- **Special relativity.** These rotation speeds, however, are at fractions of the speed of light, and so we must allow for relativistic effects. Photons from the matter moving towards us will be ‘beamed’

and the blue side will appear much brighter than the red (receding) side. Furthermore, the whole profile will be slightly red-shifted due to the transverse Doppler shift (see description of this effect in Box 11.6).

- **General relativity.** The entire radiating volume is sitting in the deep gravitational potential well

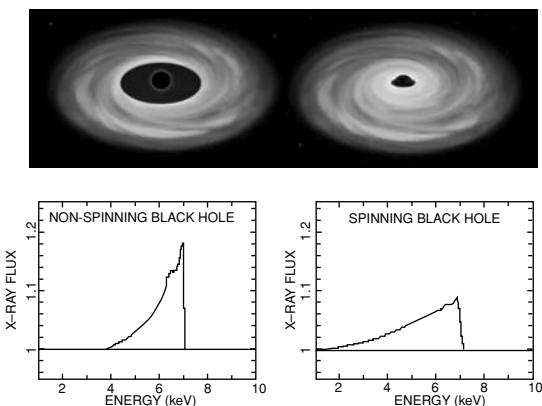


Fig 14.16 The distorted Fe K α profile is strongly affected by black hole spin. (right) A rapidly spinning (Kerr) black hole allows the inner disc to be stable down to orbits much closer to the event horizon than (left) for a non-spinning (Schwarzschild) black hole, thereby greatly changing the volume of matter where the Fe line is formed.

of the central compact object. All the photons must ‘climb out’ of this well to reach us. This introduces a *gravitational red-shift*, which can be very significant for a supermassive black hole.

When all these effects are included, then the distorted Fe K α profile of Fig. 14.13 can be well

accounted for if the disc has an inclination of about 30°. It is also possible to use the Fe K α profile to probe whether the black hole is spinning. It is an intriguing result of general relativity that the matter orbiting the disc cannot extend right down to the event horizon itself. Instead, there is an *innermost stable circular orbit* (or ISCO), inside which matter quickly plummets directly across the horizon. Even more remarkably, this ISCO is much smaller for a rapidly spinning hole than a non-spinning one, and this has the effect of greatly changing the expected shape of the Fe K α line (Fig. 14.16). The data on MCG-6-30-15 requires it to be a rapidly spinning hole. So not only are we now directly observing extreme physical effects that are expected close to black holes, but we can even start to infer both of the key properties of black holes: their mass and their spin.

14.4.4 Water masers in NGC 4258

A more detailed view of the disc surrounding an AGN has come from very high resolution radio observations (so-called VLBA, or very long baseline array, which exploits radio telescopes spread across the globe) of NGC 4258. This spiral galaxy (Fig. 14.17 (left)) is just over 20 million light years

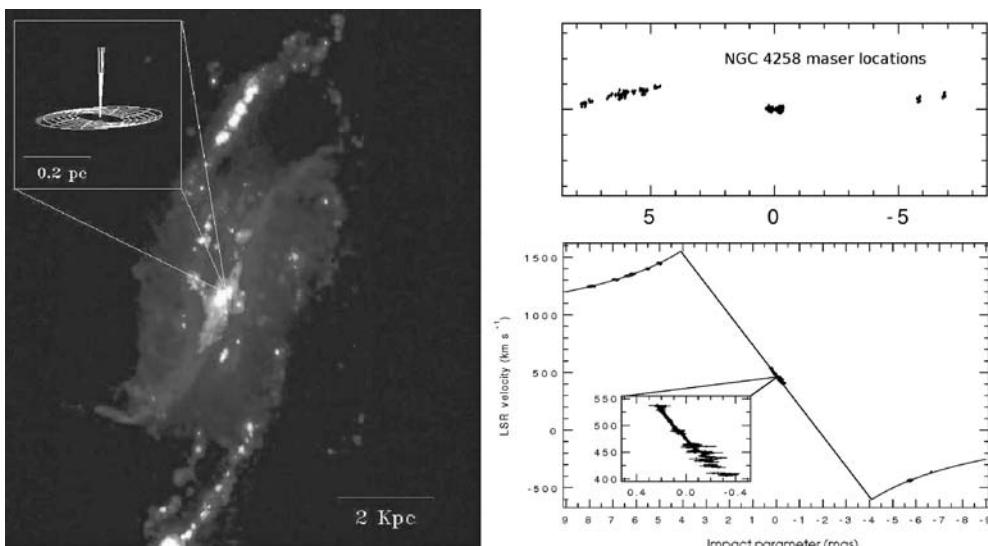


Fig 14.17 (left) Narrowband (H α) visible image of the spiral galaxy NGC 4258, whose nucleus displays radio jets and water masers (shown schematically in the inset). (right) Water maser locations (above) are extremely close (<10 milliarcseconds) to the core, and their velocities (below) fit perfectly (solid line) with a model of simple Keplerian motion about a central object of 40 million M_\odot (images courtesy of NRAO, Gerald Cecil and Holland Ford).

away and has an active nucleus with radio jets. More remarkably, Miyoshi *et al.* (1995) discovered a line of water masers very close to the core (<10 milliarcsecs away, or less than a light year at this distance), which are simply clouds of water molecules that amplify the radio emission from the core and beam it in our direction (masers are microwave equivalents of lasers, in which light is amplified by stimulated emission of radiation). Not only did Miyoshi *et al.* resolve the locations of these masers (figure 14.17 (right)), but they also measured their radial velocities. Their motion around NGC 4258 could be perfectly modelled as simply Keplerian orbits around a central object, whose mass was very accurately determined as 40 million M_{\odot} .

Chandra X-ray imaging of NGC 4258 (Colour Plate 40) shows an additional dimension to what is occurring inside this galaxy. The dominant feature is referred to as the ‘anomalous’ arm, which appears to be emanating from the core and extends across the entire galaxy, whereas the classical spiral arms in the visible are barely noticeable in X-rays. It has been suggested that this X-ray arm is a pair of X-ray jets, but it does not align with the VLA-mapped very small jets in the core. The hot material in the anomalous arm is probably the result of shock heating.

14.4.5 Stellar proper motions in the Galactic Centre

The highest spatial resolution yet obtained by direct observation applies to the nucleus of our own Galaxy, simply because of its proximity. Already mentioned in Chapter 13, the nucleus of the Milky Way is not currently ‘active’, but there is an enigmatic source, called Sgr A*, located at the exact centre that does emit at radio, IR and X-ray wavelengths, with detections being reported all the way to ultra-high-energy γ -rays. The nature of this nuclear source has been a subject of controversy for almost 40 years, but the very high obscuration (by gas and dust in the plane of our Galaxy) makes it impossible to view in the visible. To penetrate this obscuring material, it is necessary to move into the IR, but it is only since the early 1990s that it has been possible to obtain IR images from the ground with sufficient spatial resolution (Colour Plate 41) to follow the remarkable

stellar ‘dance’ that is occurring in our own galactic nucleus.

Very sophisticated adaptive-optics techniques to correct for the blurring effects of the Earth’s atmosphere have been used on both the VLT and Keck telescopes to take diffraction-limited IR images of the Galactic Centre region over the last 15 years. This stunning work, led by Reinhard Genzel (with VLT) (Genzel & Karas, 2007) and Andrea Ghez (with Keck) (Ghez *et al.*, 2005), has produced the images shown in Fig. 14.18, which have revealed how these bright, massive stars are actually moving on the sky in orbits around the central radio source, Sgr A*. One of them, star S2, has already been observed through a complete orbit, which lasted 15.2 years and is clearly eccentric. Its closest approach was observed in 2002, when it approached to only 17 light-hours from the central object. At the distance of the Galactic Centre (8.3 kpc), this corresponds to only 3 times the size of the Solar System, and S2 at that point was travelling at over 5000 km s⁻¹. With the stellar orbit and period so well defined (the semi-major axis, a , is \sim 900 AU), the mass of the unseen object coincident with Sgr A* can be easily computed from Kepler’s third law as 4.3 (\pm 0.4) million M_{\odot} . So, while it may be inactive now, the centre of our Galaxy has clearly been a place of considerable activity in the past. Furthermore, this supports the view that central supermassive black holes may be the rule in normal galaxies rather than the exception.

Much research continues into the properties of these objects populating the centre of our Galaxy. The bright cluster of stars that surround Sgr A* (Colour Plate 41), whose orbits were used to measure the central black hole mass, is itself poorly understood. As an unusual combination of both late type, evolved stars and hot, massive stars, it is thought that this mixture was produced from a cluster of young stars that was captured recently by the massive central object. Furthermore, the eventual decay of the orbit of star S2 will lead to it being tidally disrupted and accreted by the central object, likely returning it to an ‘active’ phase.

The importance of the NGC 4258 water maser observations, compared to reverberation mapping or HST spectroscopy of the region closest to the nucleus, is demonstrated in Fig. 14.19. In terms

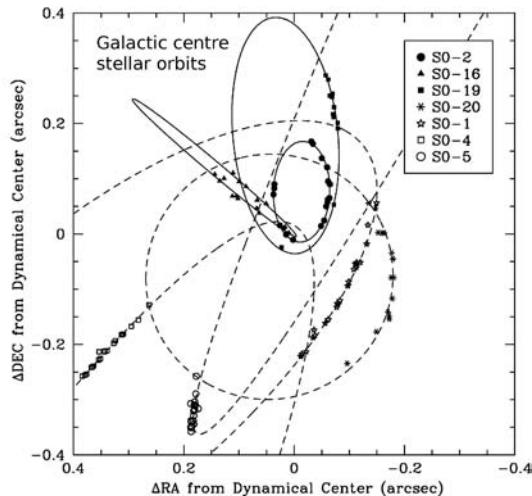
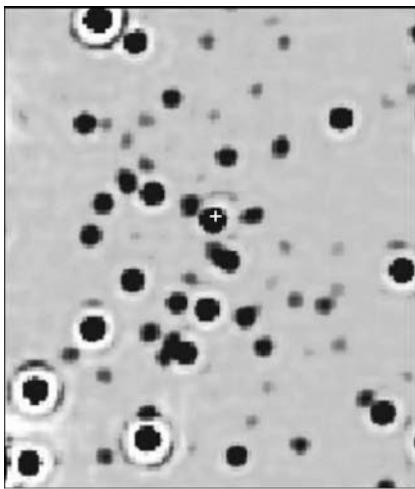


Fig 14.18 (left) Detail in the K_s band (resolution 0.06 arcsec) covering only the central $\sim 2 \times 2$ arcsec section of Colour Plate 41 showing the stars whose orbits around Sgr A* have been determined over the last 15 years are plotted (right). The location of Sgr A* is marked with an 'X' (images courtesy of ESO).

of the proximity to the event horizon, measured in Schwarzschild radii of a black hole of the measured mass, the extremely high spatial resolution achieved with VLBI takes us to within $40\,000 R_S$. More importantly, it is the inferred mass density (within that volume) that is correspondingly very high, almost 1000 times higher than the densities required to accommodate the observations described in Sections 14.4.2 and 14.4.1.

Is there now concrete proof of the existence of supermassive black holes in the nuclei of galaxies? That is, could there possibly be any alternative explanation? Attempts to avoid the conclusion that the central object *must* be a black hole have involved looking at what properties a cluster of 'dark' stars of the same total mass would have if they were confined within the tiny volume allowed by the spatial resolution of the observations. Not surprisingly, such a dense cluster of stars would suffer very close interactions and collisions, which would eventually lead to collapse of some stars into a single object, while others would be ejected from the cluster. The timescale for this to happen depends on the stellar density, as shown in Fig. 14.19, and even for NGC 4258, this timescale is comparable to the currently measured age of the Universe. However, our proximity to our Galactic Centre gives us unprecedented spatial resolution,

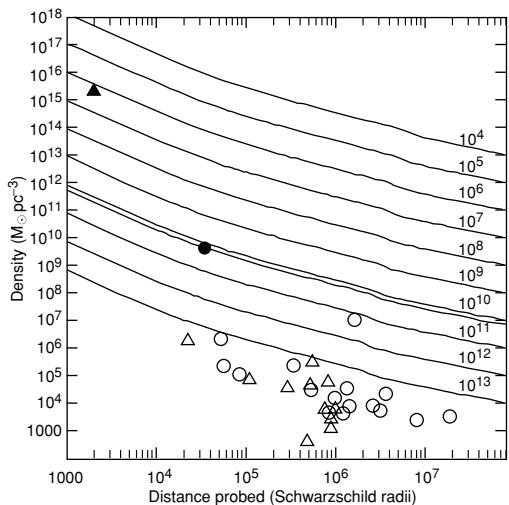
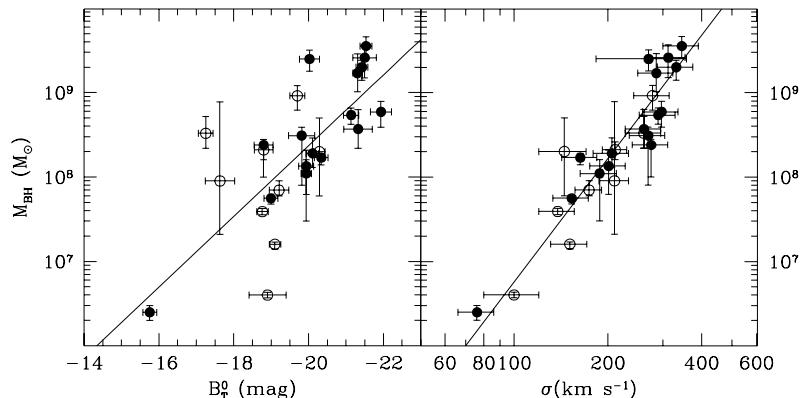


Fig 14.19 Comparison of proximity to the event horizon (in units of the Schwarzschild radius) of matter observed by different techniques. The solid triangle represents the stellar motions around the centre of the Milky Way; the solid circle represents the water maser in NGC 4258; the remaining open circles and triangles are from spatially resolved spectroscopy of the stars or discs, respectively, that are surrounding central objects in other galaxies. The inferred central density is plotted as the y axis, and the curves show the maximum lifetime in years that a cluster of stars (of the required total mass) could survive within that region. The solid curve (which runs through the solid circle) is the age of the Universe (from Ferrarese & Ford, 2005).

Fig 14.20 Relations between the dynamically measured masses of AGN central compact objects (M_{BH}) and two observed properties of galaxies: (left) the optical brightness of the central bulge of the galaxy and (right) the velocity dispersion of stars in the innermost regions. Spiral galaxies are shown as open symbols, ellipticals as filled (diagram by Ferrarese & Ford, 2005).



and so the implied density of the matter that must lie within is correspondingly much higher (a factor of 10^5 times higher than that in NGC 4258). Any attempt to explain this density with a star cluster would therefore have to explain its (astronomically) short lifetime of at most a million years. And this does not even address the question of how to hide the combined IR emission of 4 million M_{\odot} of stars! The only possibility for our Galaxy seems to be a supermassive black hole at its centre.

14.4.6 Relations with other properties

The rapid stellar motions of the star cluster in the central region of our Galaxy provide an excellent demonstration of the dramatic effects on surrounding stars and gas that can be produced by the extremely high mass densities in the nucleus. We are fortunately placed to be able to resolve these individual stellar orbits at the centre of the Milky Way, but this is not possible for any other galaxy. Instead, we measure the *velocity dispersion*, σ , of the central bulge of galaxies (the motions of stars will make the absorption lines in their spectra appear broader than that of a single star). Providing that the central bulge is sufficiently spatially resolved, then a very good correlation has been found over the last decade (see Ferrarese & Ford, 2005) between the central compact object mass M_{SMBH} and σ , as shown in Fig. 14.20 (right). This relationship can be written as $M_{\text{SMBH}} = 170 \times 10^6 \sigma^5 M_{\odot}$, where σ is measured in units of 200 km s^{-1} and has a small scatter, which means that just measuring σ alone allows

the SMBH mass to be estimated to an accuracy of about 30 per cent.

It was also first noticed by Kormendy and Richstone (1995) that there was a correlation between the central black hole mass and the visible (B-band) brightness of the surrounding stellar component. For spiral galaxies, this component would be the central bulge, whereas for ellipticals, it would be the entire galaxy. The latest version of this correlation is plotted as Fig. 14.20 (left), where it is clear that the correlation works better for ellipticals rather than spirals, but again, this provides a useful mechanism for estimating the central compact object mass.

14.5 Jets

14.5.1 Background: M87 and Cygnus A

As mentioned in Section 14.1, the first jet was discovered in early optical photographs from Lick Observatory of M87 by Heber Curtis in 1918. Figure 14.21 shows a more recent HST image of this galaxy and its jet. The radio source Cygnus A is one of the brightest radio objects in the sky, in spite of its 210-Mpc distance. It owes its radio luminosity entirely to jets which are ‘blowing’ giant radio lobes well beyond the extent of the galaxy itself, as revealed in the VLA radio map of Fig. 14.22. There is still controversy over the details of this essentially ‘classical’ QSO interpretation of Cyg A, as the great distance and optical complexity (Colour Plate 42) make this very difficult. However, the energetics are dominated by the jet, which has evacuated the

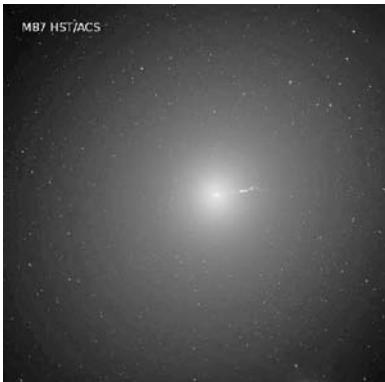


Fig 14.21 HST Advanced Camera for Surveys optical image of M87, the giant elliptical galaxy at the centre of the Virgo Cluster of galaxies, at a distance of 17 Mpc. The old stellar population of M87 appears as a diffuse glow, but at its centre is the jet, which contains large amounts of matter being ejected from close to the supermassive black hole. The many starlike images surrounding the galaxy are actually globular clusters, of which M87 has $\sim 13\,000$ (image courtesy NASA/STScI).

region around the host elliptical galaxy, producing a vast, hot X-ray cavity that is well revealed by the Chandra X-ray images. All these components are assembled schematically in Colour Plate 43, and demonstrate the basic building blocks of an AGN powered by an accreting SMBH.

14.5.2 Our nearest active galaxy: Centaurus A

This famous and beautiful galaxy of the southern skies is a remarkable object to observe at any wavelength, but it is astonishingly different in structure and detail going from optical to radio to X-rays (see Colour Plate 44). The classical (optical) image of Cen A (also known as NGC 5128) shows an elliptical galaxy, split into two by a dark dust lane which has properties normally associated with spiral galaxies, a morphology that may be the result of a galaxy merger.

Cen A's very unusual structure (with properties of both spiral and elliptical galaxies) makes it difficult to accurately determine its distance, which is currently estimated to be between 3 and 4 Mpc. This is close enough so that we can study the X-ray and radio structure of Cen A in considerable detail. Colour Plate 44 contains the individual Chandra X-ray and VLA radio maps, all to the same scale as the optical image, while the main image is a composite of all three. The extended radio lobes were first discovered in the 1950s, and the dust lane defines a plane which is approximately perpendicular to that of the radio jets. The jets are driving material into the lobes on a scale larger than the optical galaxy itself! As an outlying member of our Local Group of galaxies, Cen A is therefore the

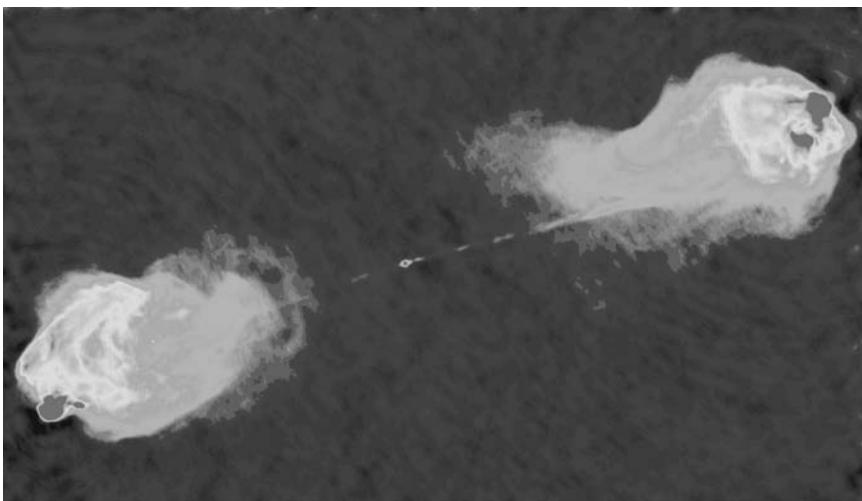


Fig 14.22 VLA radio map of the second brightest radio object in the northern sky, the giant radio lobes of Cygnus A. The high angular resolution of the VLA has revealed the (faint) emission from the jets that are powering these lobes, whose size extends well beyond the optical galaxy itself (image courtesy of NRAO/AUI).

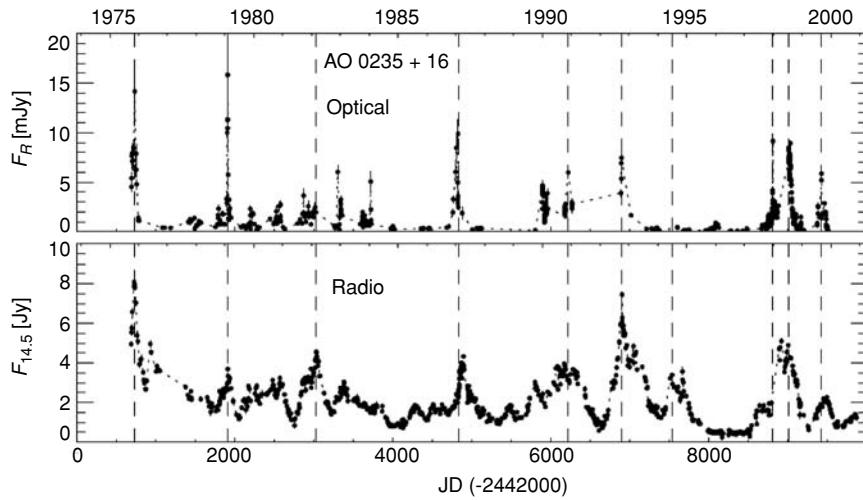


Fig 14.23 Long-term (top) optical (R-band) and (bottom) radio (14.5 GHz) flux variations in the BL Lac object, AO 0235+16. Note the variations on both long (5–10 year) and short (days) timescales (figure adapted from Raiteri et al., 2001).

nearest bright X-ray- and radio-emitting AGN (i.e. one which is currently displaying an active jet).

The composite image shows the similarity of the X-ray and radio emission from the jets, particularly in details such as knots within the jets. Earlier (Einstein) observations showed only a one-sided jet, highly collimated in the direction of the giant radio lobe, but the much more sensitive Chandra image clearly displays evidence for a ‘counter-jet’, as is seen in the radio. This asymmetry could be due to higher levels of absorbing material between us and the receding jet or an oscillation for some reason between the two directions. Nevertheless, the collimation mechanism must be stable on long timescales given the enormous length of the jet, and this is almost certainly linked to the spin axis of the SMBH which aligns the accretion disc as in Colour Plate 43 and Colour Plate 39.

14.5.3 Detailed jet structure

These jets consist of particles (mostly electrons and protons) moving at relativistic speeds, as is demonstrated in the sequence of HST images shown in Fig. 14.24. Covering an apparent distance of 24 light years in just 4 years gives a *superluminal* speed of $6c$. This is, however, just an illusion, which was encountered in Chapter 12 with the micro-QSOs and explained in Box 12.1. The

Superluminal Motion in the M87 Jet

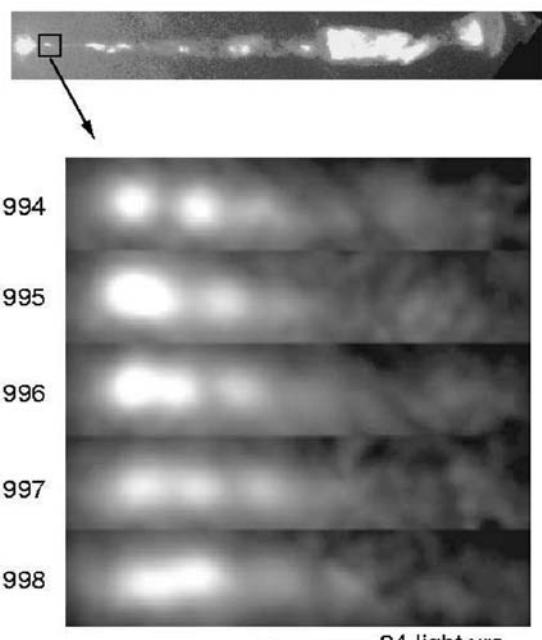


Fig 14.24 (top) The M87 jet as observed by the HST Faint Object Camera. The boxed region is expanded in the lower panels to show a sequence of FOC images taken during the interval 1994–1998. The leftmost knot visible in 1994 has expanded outward, to the right, in the subsequent 4 years, apparently moving through the length of the scale bar (24 light years). Hence its superluminal speed of $6c$ (image courtesy John Biretta, STScI).

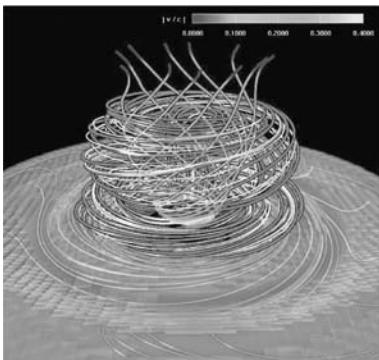


Fig 14.25 Schematic of an X-ray/radio jet in an AGN (or microquasar), in which the spin of the central compact object leads to the 'winding' up of the disc's magnetic field. This field then provides a natural avenue for the acceleration and collimation of charged particles (figure adapted from Kato et al., 2004).

phenomenon is associated with relativistic jets pointing very close to our line of sight.

The closest and brightest QSO, 3C273, gives us the most detailed look at the workings of a relativistic jet. It has been well observed by HST, Chandra and radio interferometry (Colour Plate 49). The emission is due to synchrotron radiation of relativistic electrons spiralling in a tightly wound magnetic field. X-rays require the highest energy electrons and so are brighter close to the QSO core, whereas the radio emission is brightest at the outer hot spot, where the jet encounters the intracluster hot gas and terminates.

How are these powerful jets created? It is conjectured that the combination of gas pressure from the disc, the intense radiation pressure near the event horizon (due to the extremely high temperature of the accreting material) and the rotational energy of the black hole can produce a naturally collimated jet ejected along the spin axis of the hole. The detailed physics of this process is complex and far from being understood. The most popular models are associated with tightly wound magnetic fields along the spin axis (Fig. 14.25) that are able to account for the remarkable narrowness of the jets and the relativistic velocities that they can achieve. The magnetic fields are thought to be either threaded through the event horizon (Blandford & Znajek, 1977) or a property of the accretion disc itself (Blandford & Payne, 1982). These models

have been investigated with the latest generations of supercomputers and are opening up entirely new fields of research into the properties of jets and the interpretation of the very detailed radio images.

That jets can have a significant impact on surrounding matter, and even trigger huge bouts of star formation, is shown in Colour Plate 45. This composite X-ray/optical/radio image (with an illustrated explanation to the right) shows a pair of interacting galaxies, both of which have SMBH cores but only one of which is producing a radio jet. This jet emanates from the larger galaxy and then impacts the companion galaxy, subsequently deviating. The companion galaxy appears to be rotating into the path of the jet.

An additional factor is that, even at these extremely high speeds, the time for the jets to travel from the QSO core to the radio lobes is very long (a few $\times 10^5$ years), and yet the relativistic electrons should lose their energy as synchrotron radiation in less than one-third of this time. Hence these electrons have to be accelerated en route by processes which are not yet well understood.

14.5.4 BL Lac Objects

A discussion of extra-galactic jets is not complete without the enigmatic and oddly titled *BL Lac objects*. All the jet sources so far considered have been discovered through obvious jet structure as projected onto the sky at optical, radio or X-ray wavelengths. However, with a random distribution of jet orientations in space, some will obviously be pointing in our direction. What would such a system look like? The prototype of this class, BL Lac, is a bright (~ 13 mag) variable star in the constellation of Lacertae that has been known, but not understood, for 70 years! The large amplitude (more than 2 mags), irregularity and wide timescale of the variations (e.g. Fig. 14.23) are classical properties of this group.

Although bright, and relatively easy to observe with small telescopes, the physical nature of BL Lac and related objects was a mystery until recent times, largely because the optical spectrum of the star is completely featureless (see Fig. 14.4). A pure continuum spectrum is unusual and, of course, gives no direct clues to the nature of the object. Additional properties of BL Lac and similar

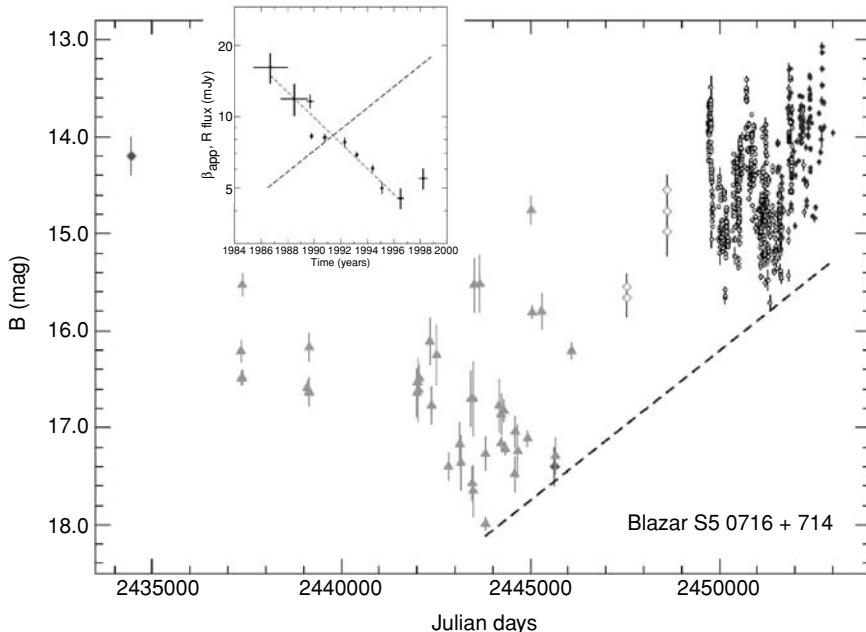


Fig 14.26 Fifty-year light curve of the blazar S5 0716+714, showing both substantial variations (>3 mags) and a recent systematic trend. Inset shows the radio measurements of the decrease in the apparent superluminal ejection speed in the jet (from $\sim 15c$ to $5c$) at the same time as the optical brightness has been increasing (figure adapted from Nesci et al., 2005).

objects include the discovery that the visible light is polarised, and they are strong radio sources. The presence of polarisation has led to their sometimes being referred to as *blazars*.

With no spectral features of any kind to measure, until the mid-1970s even the distance and hence luminosity of any member of this class were unknown. However, deep photographs of BL Lac revealed a faint fuzz around the star, which was difficult to observe spectroscopically because of the bright nucleus. Nevertheless, Joe Miller (from Lick Observatory) found that the spectrum was that of a giant elliptical galaxy at a red shift of $z = 0.07$. Several other BL Lac objects have been associated with elliptical galaxies in the same way, thus making them analogous to Seyfert galaxies, which are essentially normal spiral galaxies that have a bright nucleus. The absence of emission lines is still a puzzle that is not fully understood. The most widely accepted model of a BL Lac object is that it is an AGN in which the jet is pointing almost directly at us! Relativistic (beaming) effects enhance the apparent brightness of the continuum radiation and thereby swamp the outer emission line regions.

14.5.4.1 Precessing jets?

Since they are very bright both optically and in the radio, it has been possible to monitor a number of blazars with relatively small telescopes for some time. Indeed, archival photography has even helped with this work, as already shown in Fig. 14.23. Over the last 20 years, these programs have yielded some remarkable results, which strongly support the basic idea that blazar jets are pointed almost directly at us. Figure 14.26 shows a 50-year light curve of the blazar S5 0716+714, in which the large-amplitude optical variability is clearly evident. But VLBI radio observations have revealed superluminal expansion features (i.e. the jet) whose velocities have been decreasing over the last decade or so, as shown in the inset. At the same time, the optical brightness has been increasing. This is exactly what is to be expected if the jet is precessing slowly across our line of sight. The basic jet speed is of course constant, but if the jet axis moves from around 5° to $<1^\circ$, then the beaming factor into our line of sight will change. For a bulk Lorentz factor of ~ 12 , this change will produce exactly the optical increase that is observed.

Similar variations have been seen in the blazars OJ287 and BL Lac itself. Indeed, they appear to be periodic, with timescales of a few to ten years, and this gives a clue as to the cause of the precession. These may well be binary black-hole systems, something that is to be expected as giant elliptical galaxies interact with and accrete smaller galaxies. It is also a process that was more prevalent in the early Universe, when primordial galaxies were much closer together and interactions more frequent. In this way, larger galaxies were assembled.

A binary SMBH would lead to a precession of the relativistic jet, analogous to the precessing jets seen in galactic X-ray binaries (see Chapter 11). Calculations suggest that timescales of a few years would be expected for the precession period. Furthermore, an eccentric SMBH binary pair would produce a natural explanation for the approximately periodic outbursts seen of e.g. OJ287. With Chandra's high spatial resolution, pairs of SMBHs have now been identified, as shown in Colour Plate 46, where there are not only two SMBHs pictured in X-rays but both of them have jets which have been mapped in the radio. These jets are bent (swept back) as the galaxies in which these SMBHs are housed move through the hot, intracluster medium surrounding this system. In fact, the host galaxies are seen as NGC 1128, a ‘dumb-bell’ galaxy, but almost certainly a pair of galaxies in the process of merging.

14.6 Overall spectrum of an AGN

The proximity of 3C273 makes it the brightest QSO optically, and it can be observed over an extremely wide range of wavelengths. Its spectrum (Fig. 14.27a) is well represented by a power law of spectral index $\alpha \sim 1 (\nu^{-\alpha})$ over almost 12 orders of magnitude in frequency. Figure 14.27b shows the *spectral energy distribution* in which the spectrum is simply multiplied by ν . This gives an indication of the amount of energy radiated in equal size frequency bands and shows that 3C273 is putting out most of its energy in the X- and γ -ray regions.

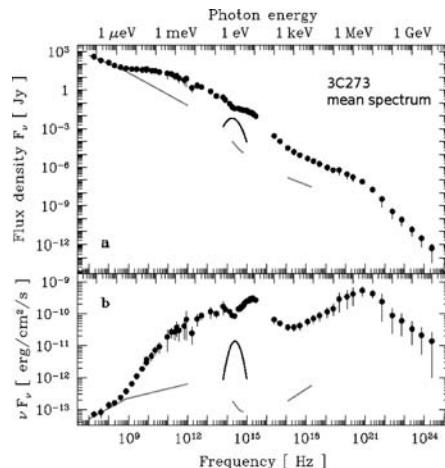


Fig 14.27 Broadband, time-averaged spectrum of the prototypical QSO, 3C273 plotted as (a) flux and (b) energy, the latter demonstrating where the peak energy output of the object can be found. The contribution from the stars of the host elliptical galaxy is indicated by the bump around $\nu \sim 10^{15}$ Hz. The lines show the emission due to the jet (image courtesy M. Türler).

There are various components contributing to this overall spectrum, which can be seen in the AGN schematic of Colour Plate 39:

- the synchrotron emitting jet, which dominates in the radio
- the thick torus, which surrounds the broad-line region and is the main absorber for high-inclination AGN, producing emission in the mid- and far-IR
- the accretion disc, responsible for the “big, blue bump” in the spectra of many AGN in the optical/UV
- X- and γ -ray emission from inverse-Compton scattering of the disc photons by high-energy electrons.

14.6.1 Accretion discs in AGN

This picture shows the exotic nature and source of violent activity in quasars and related objects. The accretion disc is an essential component. It provides a reservoir of material for accretion, helps to drive and collimate the jet and perhaps is responsible for its subsequent long-term variations. Such a disc will be much more massive than those found in galactic X-ray binaries. The direct evidence for

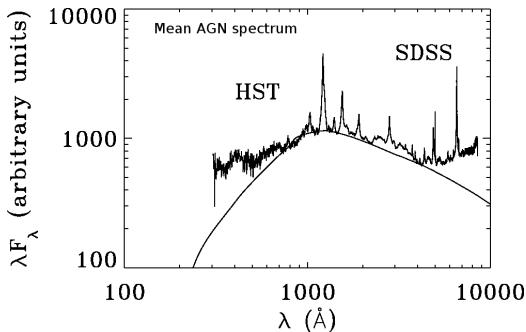


Fig 14.28 Composite spectrum of an AGN from HST and SDSS, together with (solid curve) the plot of a model accretion disc spectrum (image courtesy Omer Blaes).

the existence of such massive discs in AGN comes from examining the overall shape of the spectra from optical to X-rays, as shown in Fig. 14.28.

The radio and IR regions of AGN spectra are usually well represented by a power-law component, the expected signature of synchrotron emission by relativistic electrons spiralling in powerful magnetic fields. This is primarily associated with the jets and extended radio lobes. However, the visual and UV parts of the spectrum clearly deviate from the extension of this power law and represent a separate emitting component. This section of the spectrum peaks at around 3000 Å and is often referred to as the *big blue bump*. To account for this component, Malkan developed a model of the light output from an accretion disc based on theoretical ideas laid down in the 1970s by Jim Pringle (Cambridge), Kip Thorne (California Institute of Technology) and others. To account for the total energy output a mass accretion rate through the disc of about $10\text{--}20 M_{\odot} \text{ yr}^{-1}$ is needed. This UV hump in the spectrum provides compelling evidence in favour of the accretion disc model.

14.7 Towards the central engine: X-ray variability timescales

Variability of AGN at all wavelengths is a key property that has been known and studied for some years and is how many were originally

discovered. Several blazars were well-known variable stars long before their extra-galactic nature was realised. Variations on timescales from years down to days clearly indicated the exotic nature of AGN because this requires a prodigious amount of energy to be produced in a small volume. X-ray variability on long timescales was first discovered by early X-ray missions such as Ariel V in the 1970s for relatively nearby objects such as Cen A and NGC 4151, but this represented the limit of the technology at that time.

To probe what was going on near the central compact object required a search for variations on much shorter timescales. But how short? Variations in a bulk region of emitting gas cannot occur faster than the light travel time across that region. And this, of course, must all be outside the event horizon of the black hole. From this we expect the shortest variations to be on a timescale of about $50 M_6$ sec, where M_6 is the mass of the central object expressed in millions of solar masses. Hence, for a $10^8 M_{\odot}$ black hole, we should look for 1–2 hour variations or longer.

The discovery of X-ray emission from the whole gamut of AGN opened a new window for probing even closer to the compact object. However, the early X-ray observations of AGN revealed few clear indications of X-ray variability on short (hundreds of seconds) timescales, largely as a result of limited sensitivity. In the 1980s, it was generally thought that AGN did not vary rapidly. But the nature of X-ray astronomy operations made this a self-fulfilling statement. Observations were limited in a fundamental way for such variability studies. Satellites were launched into low orbits (encircling the Earth every ~ 100 min). The constraints of Earth occultation and charged particle regions limited observing efficiency to around 40 per cent, with no chance of long, uninterrupted observations. The first X-ray astronomy mission to breach this limit was EXOSAT. A highly eccentric 4-day orbit offered the opportunity of 90 hours *unbroken* on a single object and was the first to produce clear evidence of temporal variability due to processes occurring close to the core regions of AGN.

This capacity has been greatly augmented in the first decade of the new millennium with not one, but two, of the most powerful X-ray imaging

telescopes ever produced, NASA's *Chandra* X-ray Observatory (CXO) and ESA's X-ray Multi-Mirror Observatory (XMM). Both are still operating at the time of writing, and both are in extended orbits that allow for long, uninterrupted X-ray viewing of all types of X-ray sources. Another modern space-craft, RXTE, carries a large-area detector which is useful, but the timescale is set by a near-Earth orbit. Chapter 11 contains much that we have learnt of the long-term properties of galactic X-ray binaries as a result of the All-Sky Monitor (ASM) on RXTE. But the ASM is a small instrument, and its sensitivity limit prevents it from monitoring the fainter AGN.

14.7.1 Long-term X-ray light curves of AGN

We refer to unbroken observations within one 4-day EXOSAT orbit as short timescales to distinguish them from long timescales between separated observations of sources of weeks to months to years. The long timescales are naturally the hardest to thoroughly investigate with any space mission. Nevertheless, much has been accomplished, and the variations on long timescales in AGN seem to be mainly due to changes in the mass accretion rate. The variations are essentially the same at all X-ray energies, implying that the spectrum remains roughly constant in shape. The only exception is NGC 4151, which was constant in EXOSAT's LE (soft X-ray) telescope but highly variable in the ME (at hard X-rays). This suggests that the soft X-rays are being produced in a much larger volume such as a hot medium between the BLR clouds (see Colour Plate 39). But it was on the short timescales employing EXOSAT's unbroken looks where the major contributions have come (Fig. 14.29).

More recently, there have been long-duration observations by XMM (e.g. of the Seyfert galaxy Ark 564) (Fig. 14.30) and multiple pointings over several years by RXTE at some of the brighter AGN (Fig. 14.31).

Are there any indications of a characteristic or fastest timescale in these light curves? To look for this, the power spectra of the sources were calculated, exactly as was done for the study of oscillations in X-ray binaries (Chapter 11), and these are shown in Fig. 14.32. No single or dominant

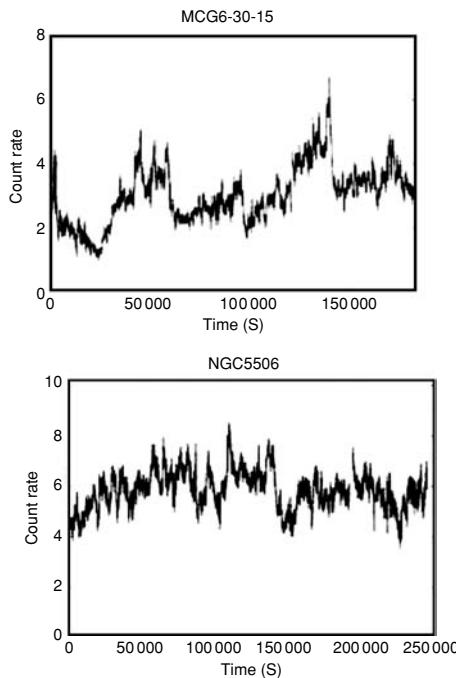


Fig 14.29 EXOSAT long looks with the hard X-ray (ME) detectors at two of the brightest and most highly variable AGN, MCG-6-30-15 (see also Section 14.4.3) and NGC 5506. There is structure in these light curves on timescales from minutes to hours (image courtesy EXOSAT Observatory, ESA).

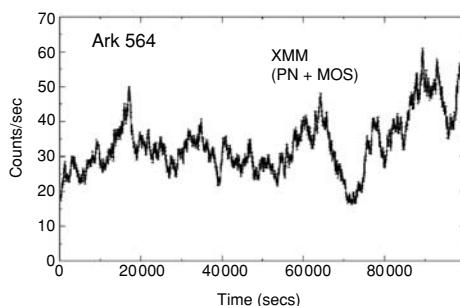


Fig 14.30 Approximately 28-hour observation of the Seyfert galaxy Ark 564 with XMM (image courtesy ESA).

timescale appears, and it is not possible to constrain the black hole mass straightforwardly. However, these power spectra do show a flattening at longer timescales (lower frequencies), with 'knees' in this plot that occur from a few times 10^{-7} Hz (a period of ~ 20 days for MR 2251-178) to 10^{-3} Hz (for NGC 4051). This flattening is also visible in

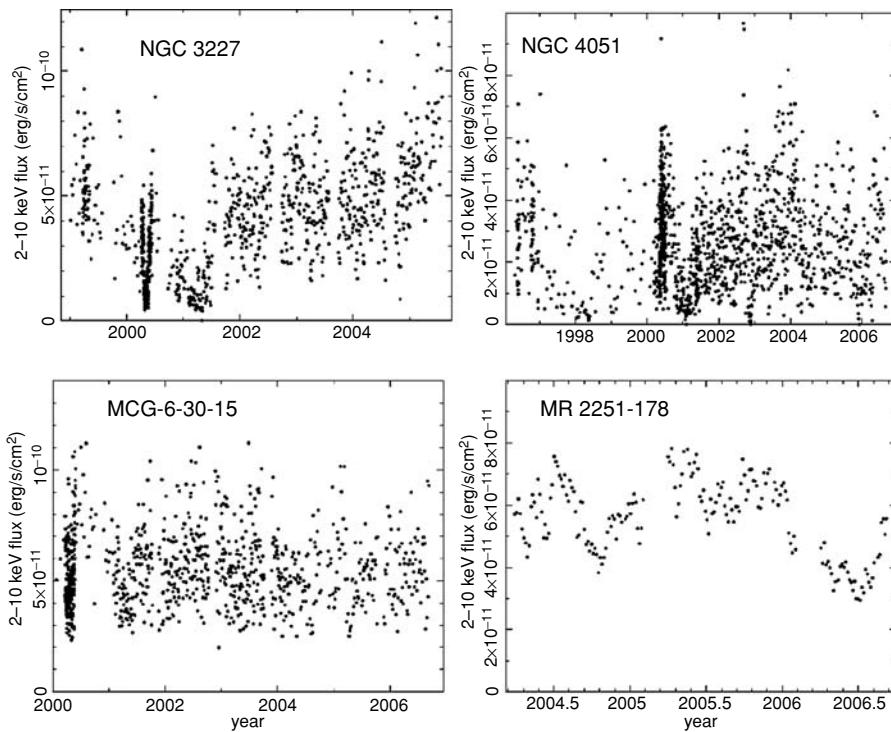


Fig 14.31 | Long-term monitoring of four AGN with RXTE. Note the presence of more densely sampled observing sequences so as to examine different variability timescales. The scatter in the data is *intrinsic* to the source. (image courtesy NASA).

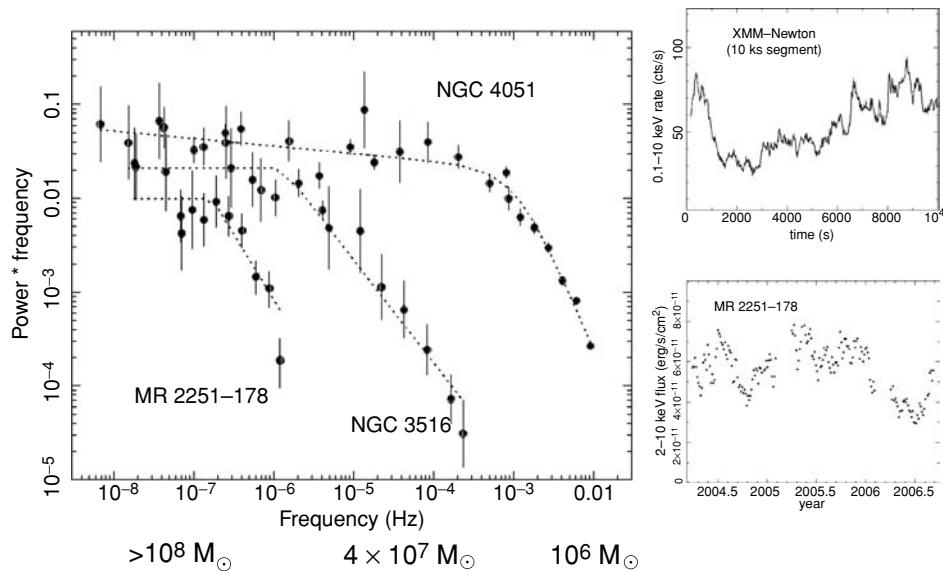


Fig 14.32 | (left) Power spectra of three AGN, (right) with samples of two of their light curves with XMM. The masses deduced from these spectra for the three AGN nuclei are given below the figure (image courtesy Phil Uttley).

Box 14.3 | A fundamental plane for AGN and XRB variability properties

McHardy *et al.* (2006) have found that the X-ray variability timescale, $t_B \propto M_{\text{BH}}/\dot{m}_{\text{Edd}}$, where $\dot{m}_{\text{Edd}} = L_{\text{bol}}/L_{\text{Edd}}$ is the mass-transfer rate onto the central BH, expressed as a fraction of that which would produce the maximum, Eddington-limited luminosity, L_{Edd} (see also Chapter 11). Consequently, it is expected that $t_B \propto M_{\text{BH}}^2/L_{\text{bol}}$, and this is exactly what they find in Fig. 14.33. This implies central object masses $M_{\text{BH}} > 10^6 M_{\odot}$ in all three cases. The timescale is due to variations in the disc itself and is not the light travel time across the regions closer to the black hole. More importantly, this relation works for galactic BH XRBs as well as AGN, a range in timescale and mass of more than 10^8 .

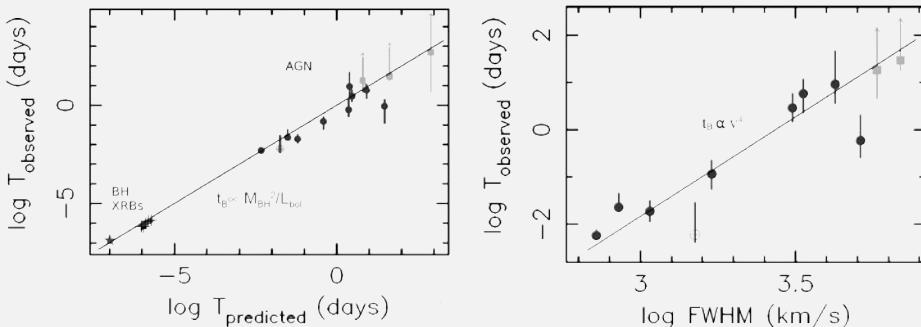


Fig 14.33 (left) Variability timescales t_B (as derived from the break in power spectra) (Fig. 14.32). Both AGN and X-ray binaries follow the relation $t_B = M_{\text{BH}}^2/L_{\text{bol}}$ over a range of more than eight orders of magnitude in t_B . (right) Variable t_B is also proportional to v^4 , where v is the FWHM of the BLR (images courtesy Ian McHardy).

the galactic black-hole X-ray binary Cyg X-1, and so McHardy and Uttley have used this similarity to scale the location of the ‘knee’, which represents a timescale of variability t_B , with the mass of the black hole, M_{BH} . The details are given in Box 14.3.

14.8 | Outflowing winds in AGN

There has been evidence for some time that AGN experience large mass loss in the form of winds. Figure 14.34 shows the sharp absorption features produced by multiple, outflowing regions

They also find a *fundamental plane* in AGN properties by examining the variability timescale, t_B , in conjunction with the motion of the BLR clouds. Assuming (as for NGC 4151; Section 14.4.1) that the BLR clouds are in Keplerian motion of velocity v at distance R_{BLR} from the central BH, then

$$GM_{\text{BH}}/R_{\text{BLR}} \propto v^2. \quad (14.3)$$

It is also observed that $R_{\text{BLR}} \propto L_{\text{bol}}^{0.5}$, and so

$$R_{\text{BLR}} \propto (\dot{m}_{\text{Edd}} M_{\text{BH}})^{0.5}. \quad (14.4)$$

Inserting this expression for R_{BLR} then gives

$$v^4 \propto M_{\text{BH}}/\dot{m}_{\text{Edd}} \quad (14.5)$$

or simply $v^4 \propto t_B$, which is plotted in Fig. 14.33.

of highly ionised material moving (blue-shifted) towards us with different speeds.

Broad absorption lines in the UV (in so-called BAL quasars) which have significant velocity *blue shifts* of up to $0.1c$ have now been joined by extensive studies of outflows via high-resolution X-ray spectroscopy using both XMM and CXO. Estimated mass-loss rates are in the range of a few M_{\odot} per year, so with velocities now measured up to $0.4c$, the kinetic energy in the outflow is comparable to the total radiated power of the AGN. These outflows are significant components in AGN energy budgets. The gain in sensitivity by XMM and CXO for X-ray spectroscopy with high-resolution

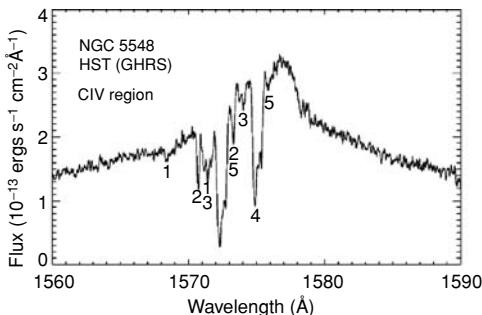


Fig 14.34 Signature of a blue-shifted AGN wind. HST GHRS spectrum of the $\text{CIV} \lambda 1550$ region in the Seyfert I galaxy NGC 5548. This AGN shows the classical broad CIV emission line, with superposed narrow CIV absorption features from a warm, highly ionised medium. There are five separate regions at different velocities producing the absorption, and each region appears twice because $\text{CIV} \lambda 1550$ is actually a doublet, or pair of lines, separated by $\sim 3 \text{ \AA}$ (figure adapted from Crenshaw *et al.*, 2003).

gratings is dramatic. Figure 14.35 shows parts of Kaspi *et al.*'s (2002) very long exposure (900 ks) on NGC 3783, then producing the best signal-to-noise and resolution spectrum of an AGN ever. A forest of absorption features are present in outflowing regions around the nucleus, and their identification with H- and He-like ions of Si and Ne indicates a high degree of ionisation which must be close to the AGN's central engine.

The detection of lines from different ionic species gives detailed information on the physical conditions prevailing where the lines are being formed. Thus the absorption is being produced in a warm, highly ionised medium (or WHIM) and is identified with the matter that is being driven from the accretion disc and surrounding areas of the AGN core. A schematic giving an overview of this interpretation (Fig. 14.36) was produced by Elvis (2000), who attempted to unify observations of AGN made at a number of wavelengths and from different intrinsic viewing angles.

The outflow from the AGN is matter driven off the accretion disc by the intense radiation

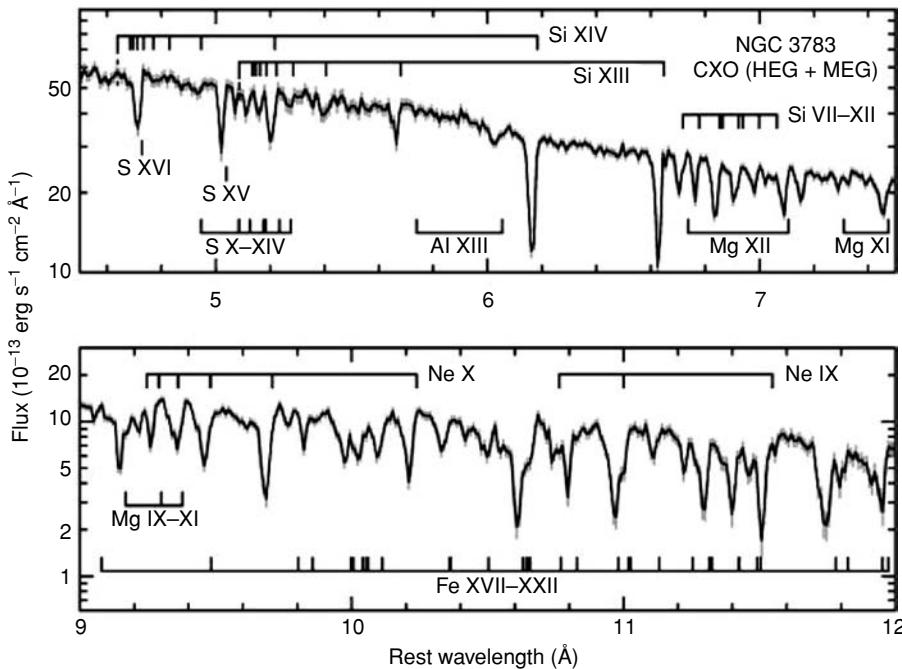


Fig 14.35 Selected regions of Chandra grating spectra around 1 and 2 keV (12 and 6 \AA , respectively) of NGC 3783, in which H- and He-like absorption lines of (top) Si and (bottom) Ne are visible. The precise laboratory wavelengths of the transitions in these series (shifted to the rest frame of the galaxy) are marked by the bars. Lower ionisation lines are also visible (figure adapted from Crenshaw *et al.*, 2003).

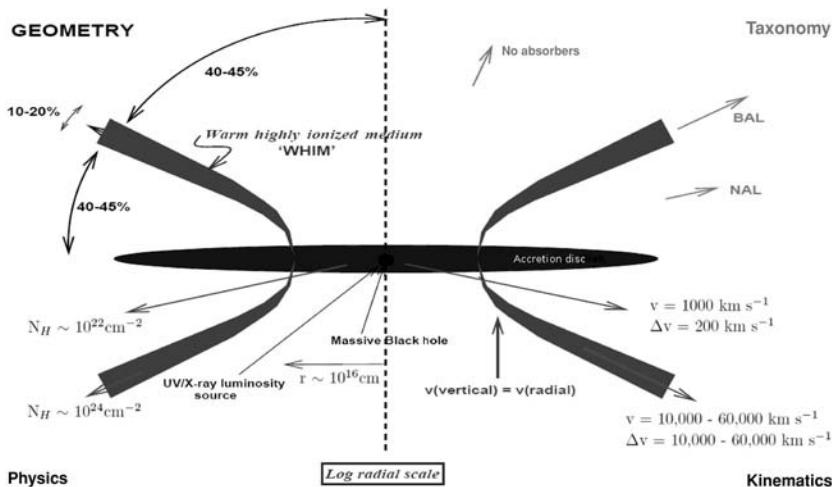


Fig 14.36 Schematic of AGN outflows showing how the same basic central engine (the accretion disc feeding the SMBH) can yield different observed properties when viewed at different inclination angles ('taxonomy' and 'geometry' quadrants). The WHIM is driven by the radiation pressure of the enormous accretion luminosity and, at appropriate viewing angles, produces the complex absorption features shown in Figs. 14.35 and 14.37. The physical properties of these regions are given in 'physics' and 'kinematics' quadrants (figure by Martin Elvis).

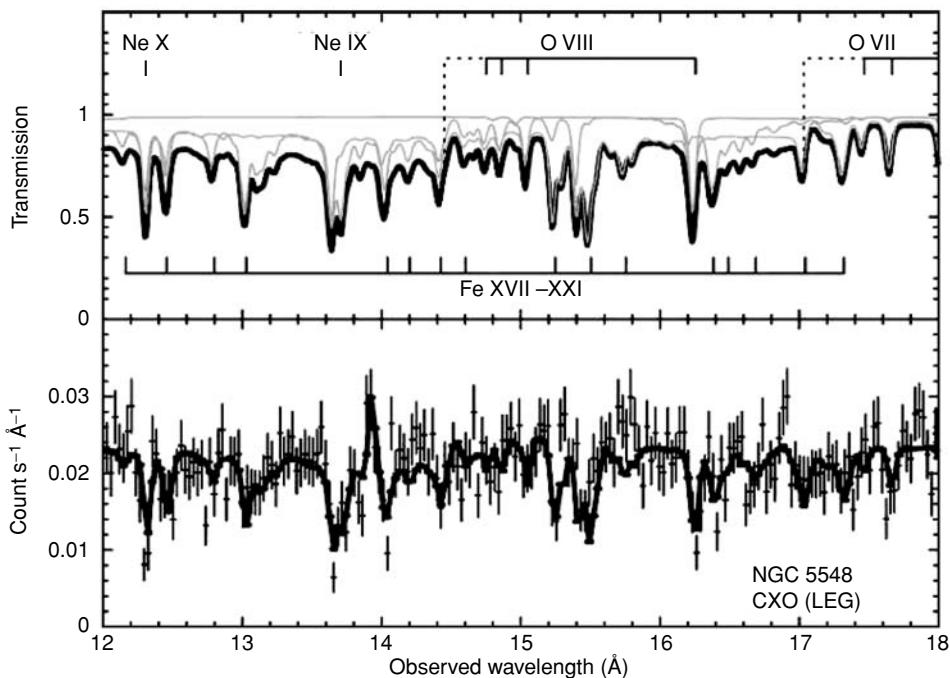


Fig 14.37 (bottom) A section of the Chandra low-energy grating spectrum of NGC 5548, where a number of absorption features are visible. The spectrum has been modelled with three different warm, absorbing components, and these are shown separately (top) together with the positions of the principal ions involved. The sum of these components is plotted in bold and also on the data itself (figure adapted from Crenshaw et al., 2003).

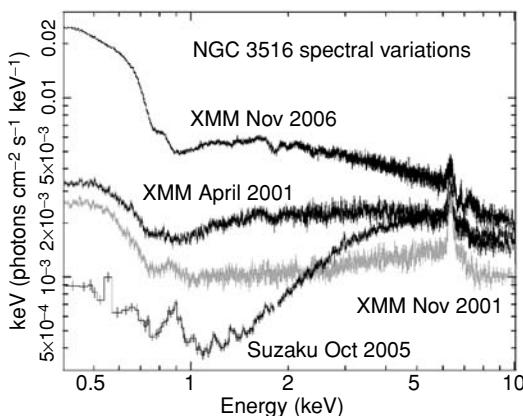


Fig 14.38 Large spectral variations in NGC 3516 observed over a number of years by XMM and Suzaku. Such changes require different column densities of the WHIM and have their greatest effect at lower energies (figure by James Reeves).

pressure of the central source. This process is expected to generate clumpiness and to be variable. At certain viewing angles, the absorption features will vary in strength and ionisation balance. Indeed, huge changes in X-ray spectral shape have been seen in e.g. NGC 3516 (Fig. 14.38) as the column density of material changes in the WHIM through which we are viewing. The WHIM will also consist of multiple components at different temperatures and hence different levels of ionisation. Such multiple components are nicely shown

in the Chandra low-energy spectrum of NGC 5548 (Fig. 14.37).

Even higher ionisation absorption features have been observed with both XMM and CXO, as shown in Fig. 14.39, where Fe^{xxv} and Fe^{xxvi} (He- and H-like ions, respectively) can be seen (and are variable in the NGC 1365 spectra). The column densities implied by these features in MCG-6-30-15 are very high ($> 10^{23} \text{ cm}^{-2}$), and these observations provide valuable constraints in interpreting the broad Fe K α line profile that was described in Section 14.4.3.

14.9 | Summary

The study of AGN has been the discovery and unmasking of massive black holes. All AGN appear to emit X-rays. The brightest were seen in early X-ray observations, and modern instruments are capable of gathering useful information from a great variety of objects and out to extremely high red shifts. Accretion powers the emission, and the X-rays allow us to probe regions closer to the black hole than we can with optical methods. Although most emission is associated with an accretion disc, spinning black holes also produce jets which are detected and studied at all wavelengths. Most of the variety of different AGN characteristics are probably due to the direction from which they are observed. Aside from the desire to know why AGN

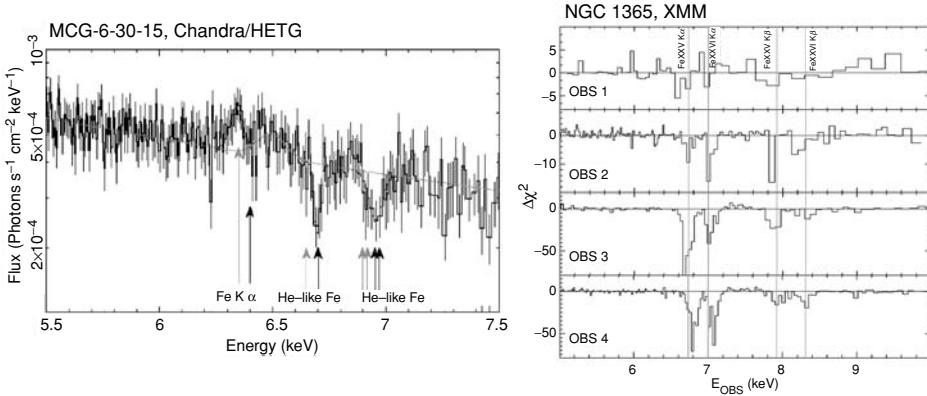


Fig 14.39 (left) A long (500 ks) Chandra HETG spectrum of the Seyfert I galaxy, MCG-6-30-15, which revealed very highly ionised Fe features (H- and He-like). (right) XMM observed similar features in the Seyfert NGC 1365 (spectra from Young et al., 2005, Risaliti et al., 2005).

appear as they do, the goal is to understand the physical processes that occur close to the event horizons and control the evolution of the Universe.

References

- Antonucci, R. R. J. & J. S. Miller (1985) Spectropolarimetry and the nature of NGC 1068, *ApJ* **297**, 621.
- Blandford, R. D. & R. L. Znajek (1977) Electromagnetic extraction of energy from Kerr black holes, *MNRAS* **179**, 433.
- Blandford, R. D. & D. G. Payne (1982) Hydromagnetic flows from accretion discs and the production of radio jets, *MNRAS* **199**, 883.
- Crenshaw, D. M., S. B. Kraemer & I. M. George (2003) Mass loss from the nuclei of active galaxies, *ARA&A* **41**, 117.
- Elvis, M. (2000) A Structure for quasars, *ApJ* **545**, 63.
- Ferrarese, L. & H. Ford (2005) Supermassive black holes in galactic nuclei: Past, present and future research, *Space Sci Rev* **116**, 523.
- Ford, H., et al. (1994) Narrowband HST images of M87: Evidence for a disk of ionised gas around a massive black hole, *ApJ* **435**, L27.
- Genzel, R. & V. Karas (2007) The Galactic Center, *IAUS* **238**, 173.
- Hez, A. M., et al. (2005) Stellar orbits around the Galactic Center black hole, *ApJ* **620**, 744.
- Kaspi, S., et al. (2002) The ionised gas and nuclear environment in NGC 3783. I. Time-averaged 900 kilosecond Chandra grating spectroscopy, *ApJ* **574**, 643.
- Kato, Y., S. Mineshige & K. Shibata (2004) Magnetohydrodynamic accretion flows: Formation of magnetic tower jet and subsequent quasi-steady state, *ApJ* **605**, 307.
- Kormendy, J. & D. Richstone (1995) Inward bound – The search for supermassive black holes in galactic nuclei, *ARA&A* **33**, 582.
- Lynden-Bell, D. (1969) Galactic nuclei as collapsed old quasars, *Nature* **223**, 690.
- McHardy, I., et al. (2006) Active galactic nuclei as scaled-up Galactic black holes, *Nature* **444**, 730.
- Mirabel, I. F. (2007) Black holes: From stars to galaxies, *IAUS* **238**, 309.
- Miyoshi, M., et al. (1995) Evidence for a black hole from high rotation velocities in a sub-parsec region of NGC4258, *Nature* **373**, 127.
- Nesci, R., et al. (2005) The long-term optical variability of the BL Lacertae object S5 0716+714: Evidence for a precessing jet, *AJ* **130**, 1466.
- Pounds, K. A., et al. (1990) X-ray reflection from cold matter in the nuclei of active galaxies, *Nature* **344**, 132.
- Raiteri, C. M., et al. (2001) Optical and radio variability of the BL Lacertae object AO 0235+16: A possible 5–6 year periodicity, *A&A* **377**, 396.
- Risaliti, G., et al. (2005) Highly ionised iron absorption lines from outflowing gas in the X-ray spectrum of NGC 1365, *ApJ* **630**, L129.
- Ross, R. R. & A. C. Fabian (1993) The effects of photoionization on X-ray reflection spectra in active galactic nuclei, *MNRAS* **261**, 74.
- Tanaka, Y., et al. (1995) Gravitationally redshifted emission implying an accretion disk and massive black hole in the active galaxy MCG-6-30-15, *Nature* **375**, 659.
- Taylor, G. B., et al. (2001) Magnetic fields in the 3C 129 cluster, *MNRAS* **326**, 2.
- Young, A. J., et al. (2005) A Chandra HETGS spectral study of the iron K bandpass in MCG-6-30-15: A narrow view of the broad iron line, *ApJ* **631**, 733.

Chapter 15

Clusters of galaxies

15.1 Introduction

On the largest scale, the distribution of matter in the Universe is uniform, but on an intermediate level, galaxies are found in gravitationally bound aggregates. These ‘groups’ and ‘clusters’ exist in sizes ranging from a few galaxies to 10 000 galaxies. The gravitational potential which binds galaxies in a cluster also binds a cloud of hot gas which fills the space between and around the galaxies. This gas, the intracluster medium (ICM), has a temperature of tens of millions of degrees. It coexists with the galaxies and, although very diffuse, is a strong source of X-ray emission.

This hot gas was discovered unexpectedly in 1971 through the analysis of X-ray observations. Modern observatories have now measured the X-ray luminosities of hundreds of galaxy clusters, and the morphology of emission from many brighter clusters has been well mapped. The shapes of the gravitational potentials of these clusters have been derived and the mass of X-ray-emitting gas determined. (The deeper the gravitational potential well, the faster the motion of the galaxies within the cluster and the greater the concentration of hot gas at the centre.) The mass of hot gas is typically 3–10 times greater than the mass derived from the visible luminosity of the galaxies.

The cluster gravitational potential which fits both X-ray and optical measurements requires the existence of a large hidden mass. Most clusters need a factor of ~ 10 times more mass than that observed in the form of galaxies and hot gas

together. This is *dark matter*, and the observation of diffuse X-rays from clusters can reveal how it is distributed.

Clusters evolve with time. Loosely bound irregular aggregates merge to become dense configurations with strong central potentials. Even on a cosmic timescale, this is a slow process, and most observed clusters show evidence of ongoing or past collisions. Many clusters also contain a central giant elliptical galaxy which is considerably brighter and more massive than any other galaxy in the cluster. These grow in size by accretion of the ICM gas and through galaxy mergers, which explains the unusually large mass and luminosity and their central position within the cluster.

These ideas and the supporting X-ray observations form the subject matter of this chapter. Of course, clusters and the effects of dark matter were first recognised in optical observations.

15.2 Properties of clusters: The optical legacy

Clusters of galaxies were recognised early as gravitationally bound entities. Photographic plates were searched and clusters catalogued. Gravitational potentials were calculated using Doppler shifts of member galaxies. Much more mass was inferred than that of the visible stars in the galaxies, which led to dark matter being recognised as the principal component.

In the catalogues of Abell (1958) and Zwicky *et al.* (1966), the membership of a cluster depends

Table 15.1 Optical properties of selected bright clusters.

Cluster	Type	Dominant galaxy	Velocity dispersion (km s ⁻¹)	Spiral fraction	Richness ^a	(z) Red shift	Distance (Mpc)
A1367	Early	None	830	0.40	117	0.0215	82
Virgo	Early	M87	660	0.55	–	0.0037	16
A400	Intermediate	cD?	420	0.42	58	0.0232	100
A2199	Intermediate	cD?	810	0.24	88	0.0305	130
Coma	Evolved	None	880	0.13	106	0.0235	100
A2256	Evolved	None	1255	0.24	88	0.0601	260
Perseus	Evolved	NGC 1275	1284	0.10	88	0.0183	77
A85	Evolved	cD	1445	–	59	0.0518	220

^a The number of galaxies not more than 2 magnitudes fainter than the third brightest member of the cluster.



Fig 15.1 An optical picture of the centre of the Coma Cluster where X-ray emission is brightest. The brightest objects are the giant elliptical galaxies (left) NGC 4889 and (right) NGC 4874 and a foreground star (with diffraction spikes). Smaller bright objects are mostly elliptical and lenticular galaxies. The field is 23' × 27' (from Palomar Observatory).

critically on how the cluster is defined, particularly in the outer regions where the density of galaxies in the plane of the sky is close to the background level. The examples to be shown in this chapter are well-studied massive clusters, and there is general agreement about their size and composition. We will refer to most clusters using their number in Abell's catalogue. The Coma Cluster (shown in Fig. 15.1), for example, is called Abell 1656 or A1656. Table 15.1 lists properties for some bright clusters.

The 'richness' of a cluster describes the number of galaxies included. The richest clusters

contain thousands of galaxies, the poorest only a few. Some clusters are called 'regular' or 'relaxed'. These are rich (at least a few hundred galaxies) and show high central concentrations and approximately spherical symmetry. They consist almost entirely of E (elliptical) and S0 (highly flattened elliptical) galaxies, systems without concentrations of dust (which shows as absorption features in photographs of the individual galaxies). The cluster diameters are in the range 1–10 Mpc. The Coma Cluster, for example, is a regular cluster.

In contrast to these, 'irregular' clusters show little or no symmetry. There is no marked

concentration towards the centre, although subclustering is often present. The membership contains all types of galaxies, including, perhaps, 50 per cent spirals. The Virgo Cluster is a familiar example of this type. Although the diameter is about the same as that of a regular cluster, the mass contained in an irregular cluster is only 10 to 30 per cent of that of a regular cluster.

15.3 Discovery of X-rays from the ICM

The first detections of X-rays from clusters were made with rocket-borne instruments launched by the U.S. Naval Research Laboratory. In April 1965, Friedman and Byram (1967) found X-rays from M87, and in March 1970, Fritz *et al.* (1971) discovered X-rays from NGC 1275. These are active galaxies located in the Virgo and Perseus clusters, respectively, and the emission was (correctly) ascribed to the individual galaxies, not to the clusters as a whole. In March 1969, Meekins *et al.* (1971) identified the Coma Cluster (which contains no active galaxies) as an X-ray source and pointed out that this emission was much stronger than that expected from the sum of all the individual (normal) galaxies within the cluster.

Data from the Uhuru satellite rapidly established clusters as an important class of diffuse X-ray sources: the Coma Cluster was discovered to be an extended source by Gursky *et al.* (1971). The Perseus Cluster was found to be extended by Forman *et al.* (1972). Kellogg *et al.* (1973), identified 16 Abell clusters as X-ray sources. All-sky surveys by Uhuru, Ariel V, and HEAO-1 soon identified over 150 clusters as X-ray sources.

In September 1975, iron line emission was discovered from the Perseus Cluster by Mitchell *et al.* (1976) using proportional counters carried by Ariel V. This established the X-ray emission mechanism as thermal, with a temperature of about 10^8 K. There was also a valuable clue to the origin of the hot gas. If the gas were primordial (created at the time of the big bang), it would consist of only hydrogen and helium. The detection of iron at roughly one half solar abundance (compared to hydrogen) meant that much of the gas had been

processed by stars and either stripped from or ejected by the galaxies to form the ICM. During the years 1979 and 1980, the Einstein telescope was pointed at about 400 clusters, and almost 300 of these were detected. The spatial extent of emission from the brighter clusters was mapped (Forman & Jones, 1982).

Discovery of the ICM was a big advance. Since the mass of the ICM is generally greater than the optical mass of the cluster, the directly observable mass in the Universe was more than doubled. After discovery and measure of some cluster X-ray luminosities, observations of spatial structure and spectra of nearer, larger clusters led to estimates of ICM mass. At this time, models assuming spherical symmetry and isothermal gas were used to calculate the amount of dark matter.

15.4 Evolution

A young cluster is irregular. The cluster potential is shallow and intracluster gas is relatively cool and unevenly distributed. Many of the galaxies contain gas and dust. As time passes, the galaxies will move in orbits around or through the cluster centre. Some will collide with other galaxies, and the tidal forces generated in close encounters will strip gas from the galaxies. The tidal force of the general potential also will transfer gas from galaxies to the ICM. This intracluster gas, in turn, will strip gas from rapidly moving galaxies. Both the galaxies and the gas will tend to settle toward the centre. The intracluster gas will heat up, and the cluster's gravitational potential will deepen. At this point, the cluster has 'relaxed', and the galaxies contain little gas and dust. The cluster properties become those of a regular cluster. A relaxed system has reached a stationary state where the distribution of objects does not change appreciably with time.

A cluster of galaxies is expected to relax in steps. First, the general distribution of galaxies and the cluster potential will come to a quasi-final configuration. The process which brings this about is that of violent relaxation (a curious combination of words). This is a collective effect in which galaxies interact with the rapidly changing potential created by their neighbours. This slows rapidly,

moving galaxies faster than individual collisions, and causes the core to collapse on a timescale of about 10^9 years for a rich cluster. The process is called violent because the change of cluster potential is rapid compared to the timescale of other processes. There is not a great deal of banging together of individual galaxies. Although the general potential at this point is now close to its equilibrium configuration, the galaxies are not.

The speed of evolution depends on the mass. More massive clusters will evolve to a regular state faster than those with lesser mass. The present regular clusters were probably born with a high enough density so that they are now farther along the evolutionary track than the present irregulars. None of the present irregulars, for example, have the mass of the Coma Cluster. The nearby irregular clusters will eventually achieve regularity but will not end as regular clusters exactly like those already observed.

The timescale for relaxation is comparable to the age of the Universe. A galaxy moving at 1000 km s^{-1} will take 10^9 years to pass through the core of a dense cluster and 10^{10} years to travel the entire extent of the cluster. The two-body relaxation time, during which collisions can change the velocity distribution appreciably, varies from about 10^9 years for heavy galaxies in the cores of regular clusters to over 2×10^{10} years for irregular clusters, which of course is older than the age of the Universe. If the cluster is ‘virialised’ (see Box 15.1), the galaxies follow trajectories determined by the gravitational potential of the cluster. Many galaxies observed in clusters do not appear to be far along the track of dynamic evolution. Essentially, these timescales are so long that all clusters are still evolving.

It is important to remember that when the cluster potential changes, this is caused by a change in the distribution of the dark matter. To describe cluster evolution in terms of galaxy motion and interaction implies that the dark matter is associated with the individual galaxies, as indeed at least some of it is.

15.5 Groups and clusters

Large concentrations of galaxies are clusters, smaller concentrations are groups. Our Galaxy is

Box 15.1

The virial theorem and dark matter

The *virial theorem* is applicable to any physical system of particles bounded in both space and velocity. The system can be gas molecules in a box or gravitationally bound galaxies. The theorem states that the time-averaged kinetic energy of the particles is equal to half the average potential energy. Thus average velocity is related to average gravitational potential, which is proportional to total mass.

The velocities of galaxies are determined by measuring the Doppler shifts, which yield only the components in the line of sight. If the motion of the galaxies is radial, with each passing through the cluster centre, then measured velocities close to the centre are close to the actual velocities. If the velocity vectors are isotropic in direction, then the actual velocity dispersion is greater than that measured. Both the general shape of the potential and the trajectories of the galaxies must be assumed before calculating the mass. The result is not overly sensitive to these assumptions.

For clusters of galaxies, the virial mass was always greater than the mass determined from the luminosity of the stars within the galaxies – the *optical mass*. This mass discrepancy was a factor of 30–100 for most clusters. The conclusion was clear that there was more mass present as ‘dark matter’ than is inferred from the optical luminosity alone.

Reference is sometimes made to the ‘virial radius’ of a cluster. Inside this radius a galaxy is considered to be a full member of the cluster. There may be in-falling galaxies outside this radius, but they do not have much influence on cluster dynamics. The virial radius of a cluster is sometimes taken as r_{200} , the radius for which the average interior density is $200 \times \rho_c$, the critical density (see Box 15.4)

a member of a group. The nearest neighbours of the Milky Way are the Magellanic Clouds at distances of 55 and 60 kpc. Then, a bit farther away, there are about 10 other dwarf irregular and dwarf elliptical galaxies. These form a swarm around a large spiral galaxy (the Milky Way). Together with M31 and its attendant smaller systems, we form

Table 15.2 Properties of groups and clusters.

Property	Groups	Clusters
Richness	2–30 galaxies	30–300 galaxies
Radial velocity dispersion	100–500 km s ⁻¹	400–1400 km s ⁻¹
Mass	$10^{12.5}$ – 10^{14} M _⊙	10^{14} – $10^{15.5}$ M _⊙
T _X	<2 keV	2–14 keV
L _X	< 10^{43} ergs s ⁻¹	10^{43} – 10^{45} ergs s ⁻¹
Fraction of galaxies within	~55%	~5%

Note. From Vrtilek (2007).

the Local Group, all contained within a volume of radius 1 Mpc. If viewed from a great distance, our group would appear to consist of only the two large spiral galaxies – a binary group. (Figure 13.1 is a scale drawing of the Local Group.)

The nearer bright spirals are also found in groups. De Vaucouleurs (1965) has catalogued 54 groups at distances less than 10 Mpc. The nearest of these is the Sculptor Group, a loose association of seven spiral galaxies with NGC 253, an edge-on system, being the largest. Members of this group subtend an angle of 20°, are at an average distance of 2.8 Mpc and are only visible from the Southern Hemisphere. Another nearby group is centred on the galaxies M81 and M82 in Ursa Major at an average distance of 3.6 Mpc. The boundary of this group may extend south towards the galactic plane to include the obscured galaxies Maffei 1 and 2 and about 30 smaller systems. And so in this manner, groups can be listed containing familiar galaxies extending out to the Virgo Cluster, the nearest large cluster of galaxies.

The Virgo Cluster is centred ≈16 Mpc from us and subtends an angle of 12°. It contains at least 100 fairly massive galaxies and a total of perhaps 1000 galaxies. The largest galaxy is M87, a giant elliptical and a strong source of radio emission (Virgo A). About 40 per cent of all clusters contain dominant galaxies such as this, and these play an important role in the evolution of the cluster.

A group is a smaller collection of galaxies than a cluster. The gravitational potential is not as deep, and the X-ray-emitting gas is cooler. Table 15.2 lists some differences. Groups are diverse. Some have many spiral galaxies and little gas. Others have older ellipticals and an ICM

distribution quite like a small relaxed cluster. Groups also merge into clusters with time. It is common to observe a group falling into a larger cluster as e.g. the Coma Cluster. Since most galaxies are born in and evolve in a group environment, the study of groups can lead to insights about galaxy formation. For example, collisions between individual galaxies are expected to be more probable in a group than in a cluster. Early work on X-ray emission from galaxy groups is summarised by Mulchaey (2000).

15.6 X-ray observations of ICM temperature and morphology

Instruments on most spacecraft have measured spectra of the brighter clusters. These measurements show that the diffuse X-ray emission is without doubt thermal radiation from hot gas. The setting is favourable for spectral measurements. Although the source is physically very large, the gas is thin and is completely transparent to its own radiation. No radiation is self-absorbed. Also, all elements of the gas are in ionisation equilibrium. The kinetic energy of the positive ions, the kinetic energy of the electrons and the distribution of ion states are all described by the same temperature in any small region of the gas.

The temperature of the gas can be deduced from the shape of the continuum, which, at temperatures above a few kilo-electron-volts, is almost all from bremsstrahlung generated by the collisions of free electrons with positive ions. If the gas is of normal composition (similar to that of

Box 15.2 The self-gravitating isothermal sphere: The beta model

This model is commonly used to fit the gravitational potential of clusters. It describes the dark matter which comprises almost all the mass of the system. The nature of this dark matter is unknown, but it must be distributed throughout the cluster in the manner approximated by the model. It is assumed to behave as a gas in hydrostatic equilibrium, with thermal pressure of expansion balanced by the inwardly directed force of gravity due to the mass of the material.

Two parameters completely characterise the distribution of this material, and consequently the potential: the central density, ρ_0 , and the ‘core radius’, a . The dependence of mass density, ρ , on distance from the centre, r , has been calculated (King, 1966) and in the central region is

$$\rho = \rho_0[1 + (r/a)^2]^{-3/2},$$

and the two-dimensional projected profile (which is observed) is

$$\sigma = \sigma_0[1 + (r/a)^2]^{-1},$$

where σ_0 is the projected density at the cluster centre.

The X-ray-emitting gas floats in but does not contribute appreciably to this potential. The X-ray emission is proportional to the square of the gas density, which emphasises the densest part, usually the central region. The observed surface brightness, S , can be described by the form

$$S = S_0[1 + (r/a)^2]^{-3\beta+1/2},$$

where β is the ratio of energy per unit mass in galaxies to energy per unit mass in the gas. For most clusters, β is in the range 0.5–1.0, indicating that as radius increases, the density of gas decreases less rapidly than the density of dark matter. At the core radius, where $r = a$, the X-ray surface brightness generally drops to 10–30 per cent of its value at the centre. Figure 15.2 illustrates the excellent fit obtained with this model to X-ray data.

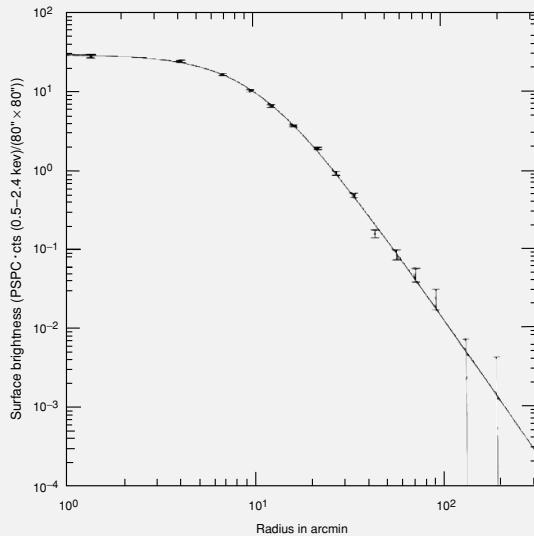


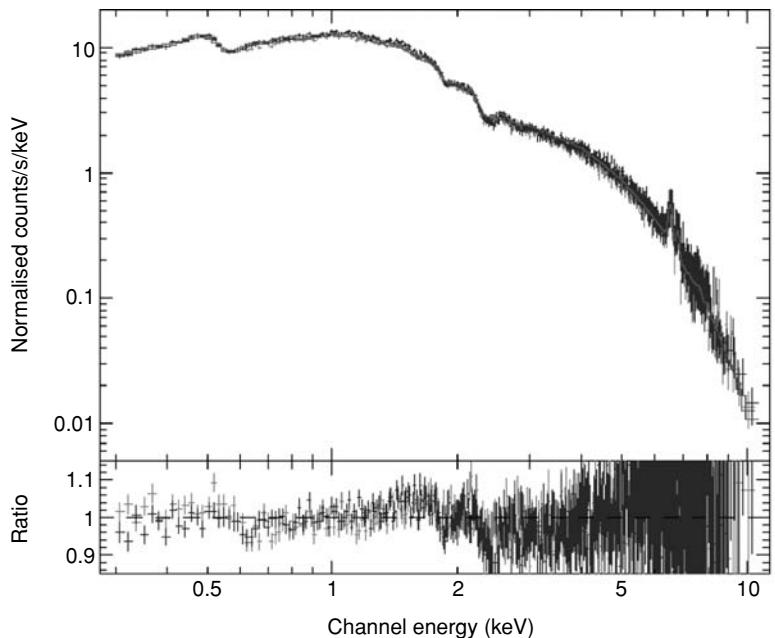
Fig 15.2 Surface brightness of the Coma Cluster as measured by the ROSAT PSPC detector (0.5–2.4 keV). The last non-zero data point is at a radius of 1.5°. The solid curve is a beta-model fit to the data. Parameters of this model give the radial distribution of dark matter in the cluster (from Briel et al., 1992).

the Sun), there are characteristic ionic lines in the spectrum. If the gas is in equilibrium, a given ion can only exist in a narrow temperature range. The presence of that particular ion thus indicates the temperature. Chapter 2 explains how continuum shape and spectral lines are determined by gas temperature.

The imaging telescopes of Einstein, ROSAT and ASCA produced maps of surface brightness. The

telescopes of Chandra and XMM produced maps of surface brightness and temperature. When photon collection was modest, as for early observations and distant clusters, mass and gravitational potential were derived using an average temperature for the ICM and assuming spherical symmetry. The model used was the beta model described in Box 15.2. The average X-ray spectrum of many clusters was well fitted using a thermal spectrum

Fig 15.3 Spectrum from the centre of the Coma Cluster measured by the XMM EPIC detectors. Spectrum is a thermal continuum with Fe emission lines at ≈ 6.7 keV. (bottom) Residuals between data and the best-fit model with $kT = 8.25$ keV (from Arnaud et al., 2001).



with a single temperature varying from 2 to 8 keV (20–90 million degrees). A recent example of this technique is shown in Fig. 15.3, where the spectrum of part of the Coma Cluster is fit with a bremsstrahlung continuum with temperature of 8.2 keV (95 million degrees). The energy of the iron line indicates a mixture of Fe xxv and Fe xxvi, as expected at this temperature.

Many clusters, however, had average spectra which required an additional low temperature component. Usually, there was also strong soft emission surrounding a large galaxy at the centre of these clusters. A high-density, low-temperature gas surrounding a central galaxy came to be called a *cooling flow*, which will be discussed in the next sections.

Except for cooling-flow clusters, single temperature fits and the beta model were used for cluster analysis until the more sensitive telescopes on Chandra and XMM showed that the diffuse gas was usually not isothermal. Relaxed clusters generally have cool cores, hot midsections, and decreasing temperature in the outer parts. Irregular clusters are evolving and have complex temperature and density profiles. Structure can be irregular due to collisions during merging or to activity in individual galaxies.

Cluster mass can sometimes be derived from gravitational bending of light from background objects and from X-ray observations using a model for the distribution and temperature of the hot gas. The X-ray observations have been most valuable in the determination of the shape of cluster potentials, often indicating structure that is not evident in the optical data. Consider the following examples.

15.7 The Coma Cluster: Rich, regular and almost relaxed

The Coma Cluster is the nearest very rich cluster and one of the brightest extragalactic X-ray sources. The diffuse X-ray source subtends an angle of $\approx 3^\circ$ in the northern sky. It has been extensively observed by all X-ray missions and was often used as a typical example of a relaxed, regular cluster. The surface brightness was fit with an isothermal beta model. Figure 15.2 shows the ROSAT result. Spherical symmetry was assumed, and two regions of obvious enhancement were excluded from the analysis. Total mass was found to be $\approx 10^{15} M_\odot$, and ≈ 20 per cent of this was the

Plate 1 Scientists and engineers with a recovered rocket payload. This was carried by a Terrier-Sandhawk rocket launched by Sandia Corp., 12 May 1970, from the Kauai Test Range. The payload reached an altitude of 306 km and spent 6 minutes above the atmosphere detecting X-rays. After re-entry, it was recovered from the ocean. White paint on the nose cone has been charred black by heating during ascent. The proportional counter slat collimator is pointed to the upper right and is intact. Left to right are: F. Seward, J. Harri, A. Iantuono, P. Stoering, T. Erven, C. Cornell, J. Zickuhr, R. Hill, and R. Grader. All are smiling because everything has worked perfectly. (Courtesy of LLL and Sandia Laboratories.)



HEAO A-1 ALL-SKY X-RAY CATALOG NAVAL RESEARCH LABORATORY

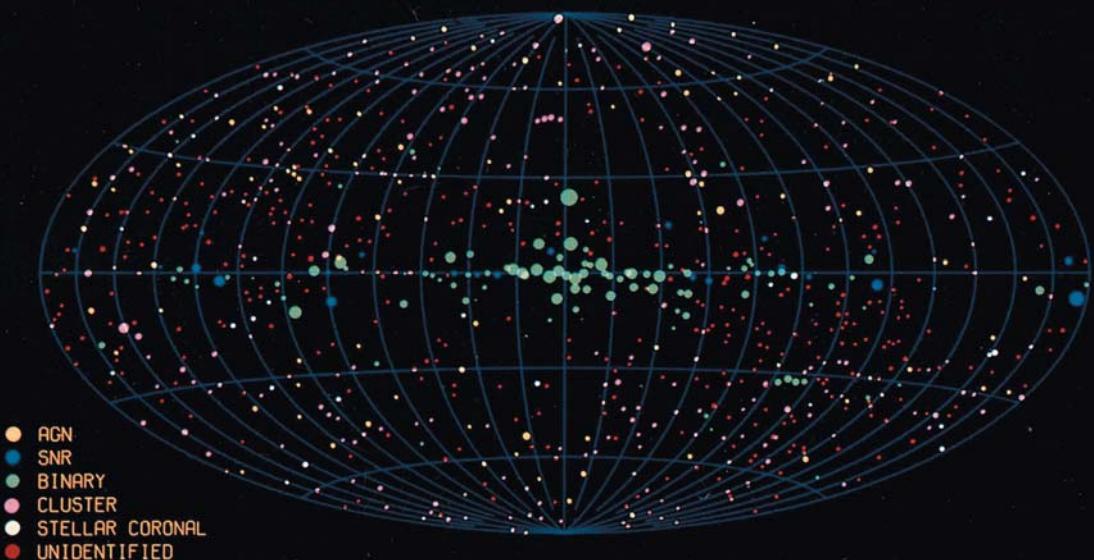
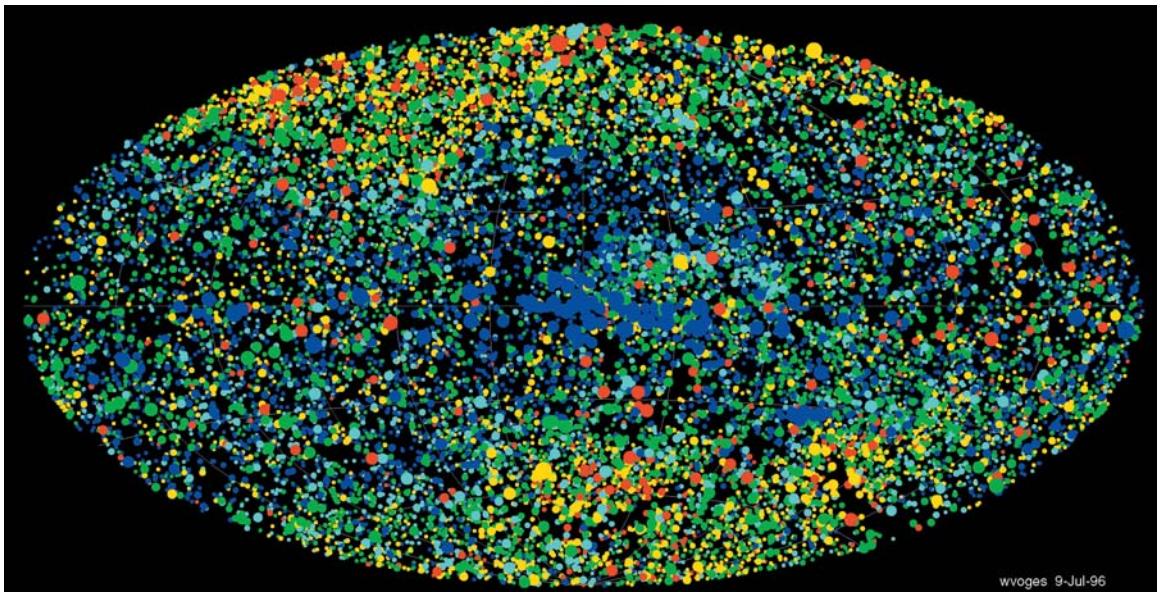


Plate 2 Map showing X-ray sources from the HEAO-1 all-sky survey in the energy band 0.5–25 keV. Size of the dot shows the brightness of the source. Colours indicate type of source. (from Wood *et al.*, 1984).



wvoges 9-Jul-96

Plate 3 Map showing X-ray sources from the ROSAT all-sky survey in the energy band 0.1–2.4 keV. Size of the dot shows the brightness of the source. Colours indicate energy – going from red (soft) to blue (hard). Absorption in the Interstellar Medium (ISM) removes all soft X-rays from sources in the galactic plane especially towards the galactic centre. (from Voges et al., 1999).



Plate 4 Postage stamp issued to commemorate the 100th anniversary of the discovery of X-rays.

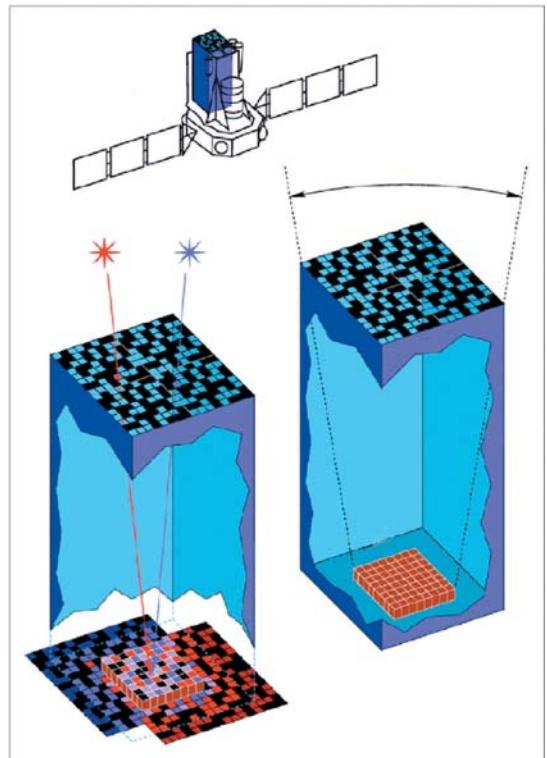


Plate 5 (1) The coded mask detector, IBIS, on the Integral spacecraft is shown at the top. (2) At lower right, the field of view is shown to be determined by the detector and mask dimensions. (3) At lower left, two sources, red and blue, cast different shadows of the mask on the detector. IBIS images X-rays and γ rays in the range 15 keV to 10 MeV with spatial resolution of 12' (ESA/IAS/CEA 2009).

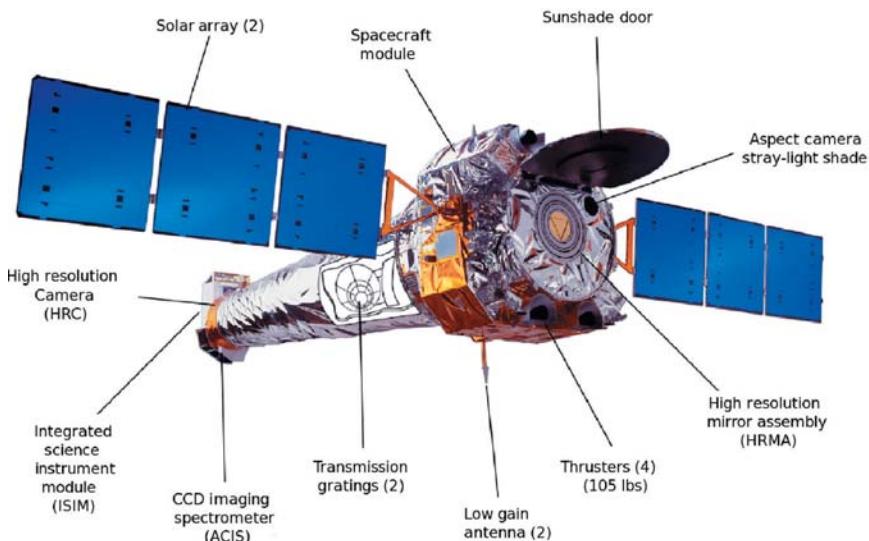


Plate 6 The Chandra spacecraft with important features labelled (NASA/CXC/NGST 2010).

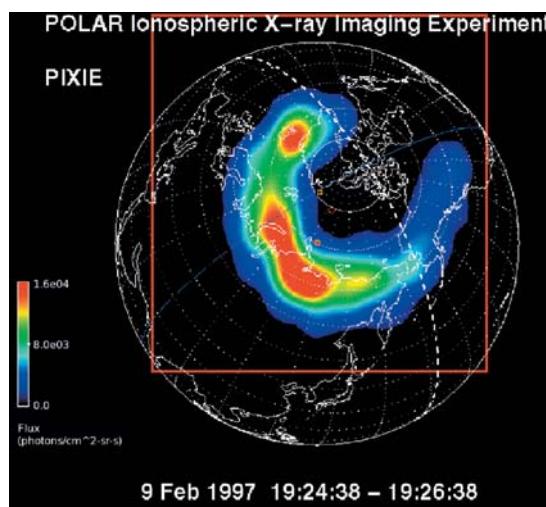


Plate 7 A bright X-ray aurora over China recorded by the PIXIE instrument on board the POLAR spacecraft. The energy range is 2–10 keV, and exposure was 2 min. The large red box shows limits of field of view. The white dashed line is the day-night terminator, and the maximum emission is on the night side. Coordinates shown are magnetic latitude and longitude. The small red square indicates the north geographic pole (NASA-GGS/ LMMSC 2008).

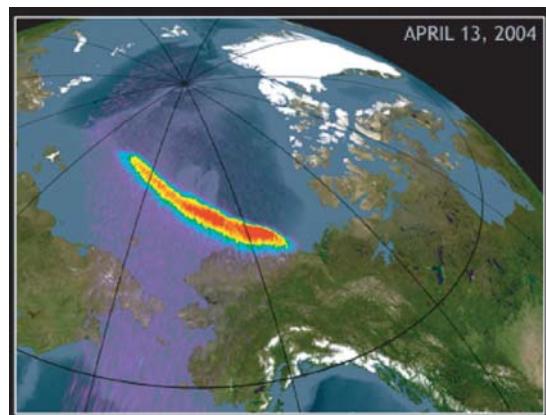


Plate 8 Auroral X-ray emission over Alaska recorded by the Chandra HRC. Energy range is 0.3–10 keV. Coordinates are geographic. This emission is much fainter than that shown in Colour Plate 7 (from NASA/MSFC/ CXC/2005, Bhardwaj et al., 2007b).

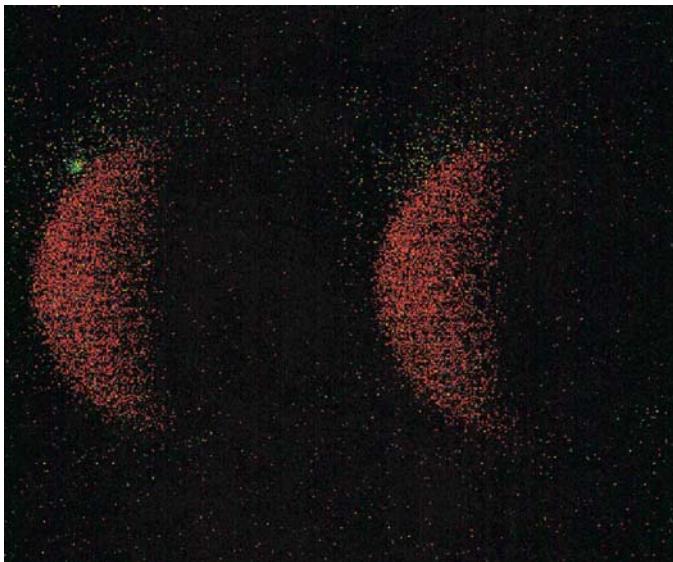


Plate 9 A very clever ROSAT observation in which the bright galactic source GX 5-1 was occulted by the moon and the dust-scattered halo is seen uncontaminated by instrumental effects. The image at left was obtained just after GX 5-1 emerged from behind the lunar limb, the image at right just before egress. X-rays from GX 5-1 (green) and the dust-scattering halo are hard because the soft component is absorbed in the ISM. Soft (red) solar X-rays scattered from the sunlit hemisphere of the Moon make the Moon easily visible to ROSAT (from Predehl *et al.*, 1992).

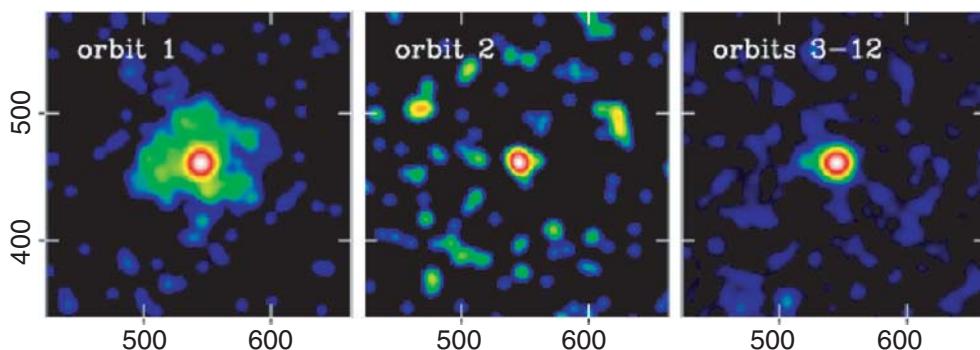


Plate 10 Images of GRB 050724 from SWIFT. The afterglow appears as a point source in all frames. Coordinates are detector pixels, each = $2.35''$. X-ray energy range is 1–5 keV. The first frame, taken ≈ 0.1 – 0.6 hours after the burst, shows the halo expanding from a radius of $\approx 50''$ to $100''$. The second frame, taken ≈ 1.7 – 2.2 hours after the burst, shows a faint ring with radius $\approx 200''$. No halo is detected in the third frame, a long integration at later times (from Vaughan *et al.*, 2006).

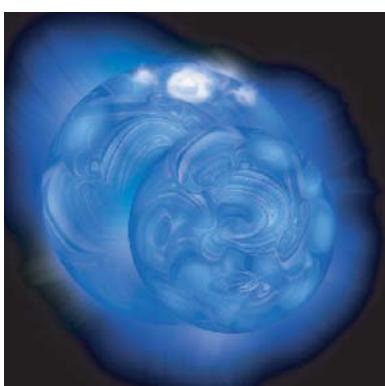


Plate 11 An artist's conception of the 44i Boo double star. Stellar size and separation are realistic, but fine-scale coronal structure is fanciful. Half of the X-rays do come from the polar region of the larger star. This contrasts with solar active regions which are equatorial (figure from NASA/CXC; M. Weiss, 2001; Brickhouse *et al.*, 2001).

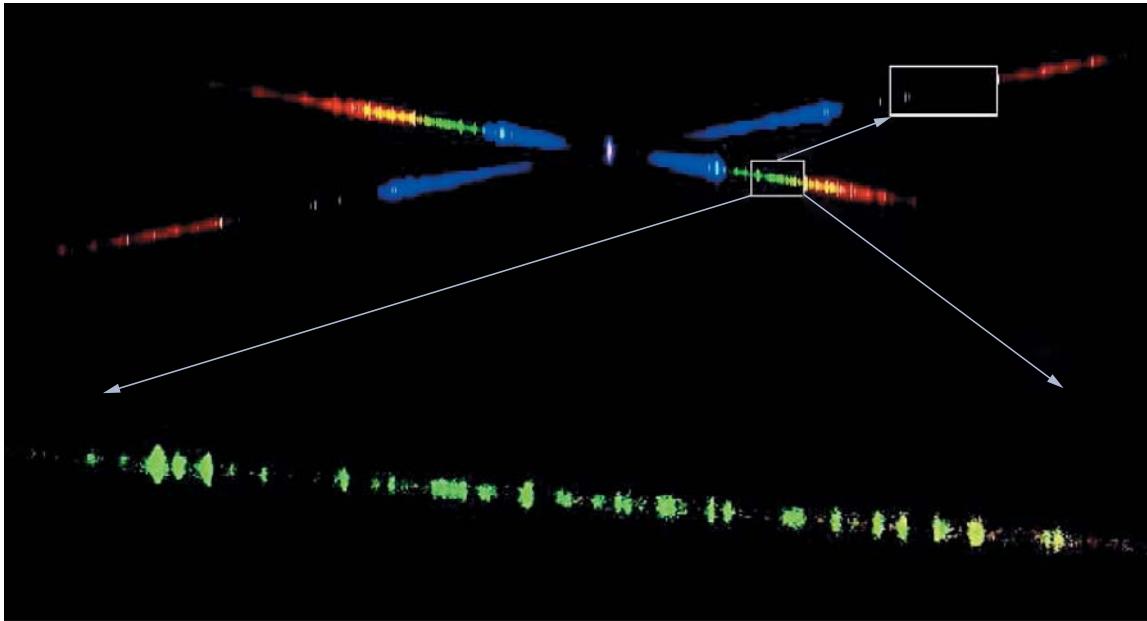


Plate 12 Three hundred kiloseconds of summed Chandra observations of Capella with the High Energy Transmission Grating. (top) Full image and the X configuration which separates spectra produced by the two sets of gratings with different periods. The MEG grating (upper left to lower right) has wider element spacing and less dispersion than the HEG. Only the first orders of both are shown, and the image has been stretched along the y axis so emission lines are visible to the eye. Colours indicate energy: red, 0.48–1.0 keV; green, 0.80–3.0 keV; blue, 1.8–6.0 keV. Boxes enclose that part of the spectrum from 9 to 12 Å which (bottom) has been extracted from the MEG and expanded (figure courtesy of Dave Huenemoerder; also see Canizares et al., 2000).

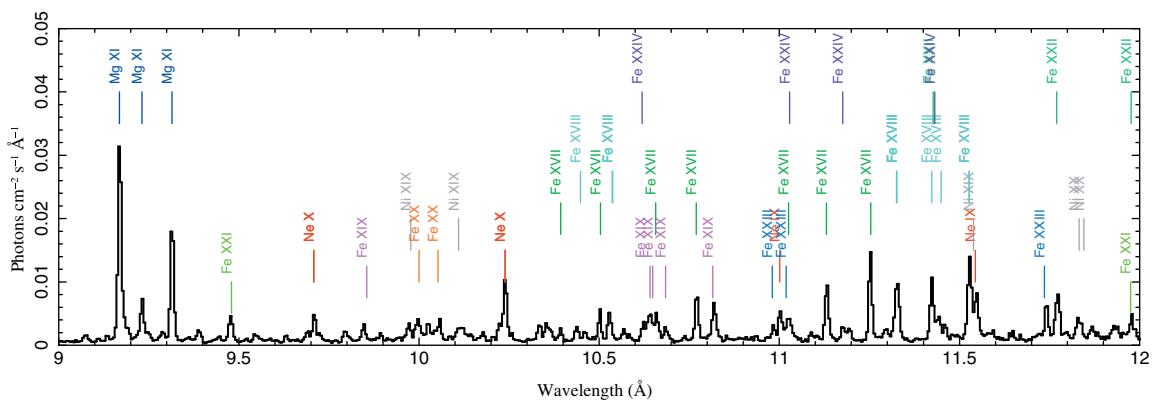


Plate 13 The spectrum extracted from the Capella data shown in the previous figure. Lines are identified with coloured labels. The bright triplet at $\approx 9.2 \text{ \AA}$ is from He-like Mg. Other lines are from various ions of Ne, Fe and Ni (figure courtesy of Dave Huenemoerder).

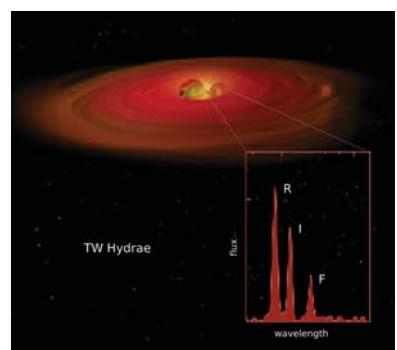


Plate 14 An artist's rendition of the star TW Hydrae and surrounding disc. An accretion stream onto the star produces X-rays. The inset shows the [O VII] spectral triplet which has a weak forbidden line. The low f/i line ratio is an indication of high density (NASA/CXC/N. Weiss 2003; Kastner et al., 2002).

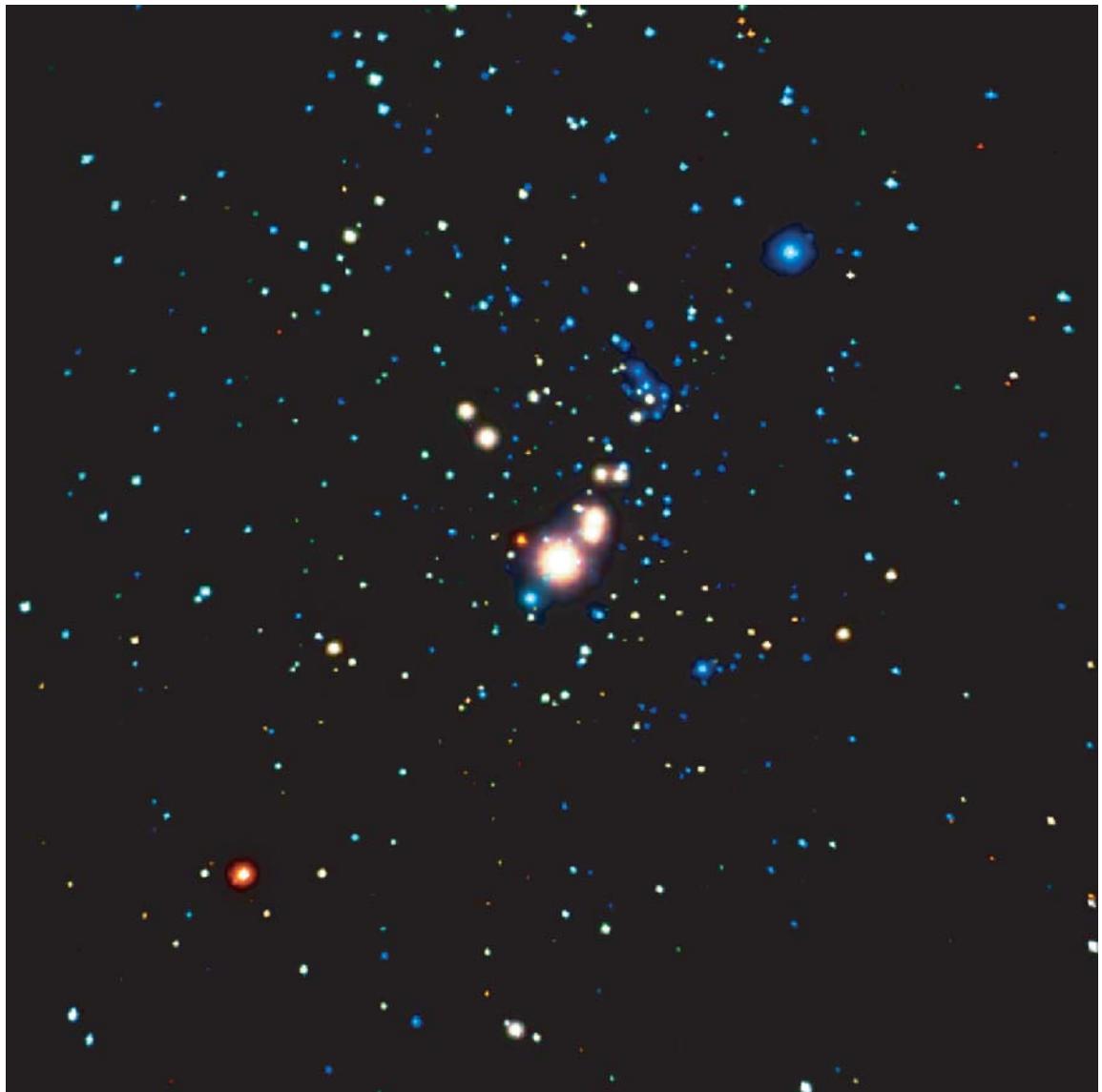


Plate 15 The Orion star-forming region imaged by Chandra ACIS in a 10-day exposure. The image is 5.5' on a side. Energy is colour coded: Red, 0.3–1.0 keV; Green, 1.0–3.0 keV; Blue, 3.0–8.0 keV. Colours are mostly due to varying absorption along the line of sight. Central star is the O7 V star θ^1 Orionis C (NASA/CXC/ Penn State/E. Feigelson/K. Getman 2005).

Plate 16 Optical and X-ray emission from η Carinae. The field size is $1.6' \times 1.3'$. Chandra observations are shown in red and yellow overlaid on a blue HST picture. The central pointlike X-ray source is at the centre of the optical nebula and is overexposed here. The bipolar optical nebula shows dust and gas thrown off by the star and is a strong infrared source. This is surrounded by faint optical wisps just inside the X-ray-emitting shell. This outer emission comes from million-degree gas heated by high-velocity material plowing into surrounding gas (X-ray NASA/CXC/ GSFC/2007; optical NASA/STScI).

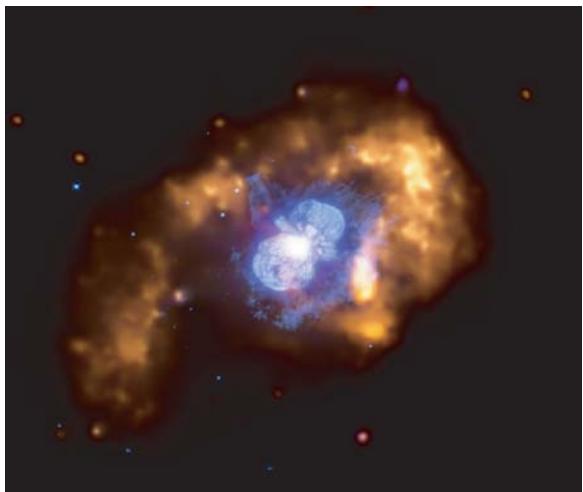


Plate 17 True colour image of the Carina Nebula, one of the brightest emission regions in the southern Milky Way, from the Curtis-Schmidt telescope at CTIO. The region shown is $3^\circ \times 4^\circ$ in extent and matches Fig. 7.10, which shows the X-ray emission overlaid. Diffuse emission is brightest above a V-shaped dust lane. Star η Car is the bright object at the lower left, and the Tr 14 O association is the bright object at the upper right of this diffuse region which fills the V (from NOAO/AURA/NSF 1975).



Plate 18 A 16-hour Chandra ACIS image of the O-star cluster Trumpler 14. This may be compared with the central part of Colour Plate 17, an optical image of the Carina Nebula. Colours indicate X-ray energy: red shows 0.5–2 keV and blue shows 2–7 keV. Data are from six ACIS chips each with an $8'' \times 8''$ field. The two lower chips are considerably off-axis here, and images of pointlike sources are not sharp. X-rays from O stars in Tr 14 form the central bright concentration. The right part of the field is shadowed by a dust lane (courtesy Lisa Townsley; Townsley et al., 2005).

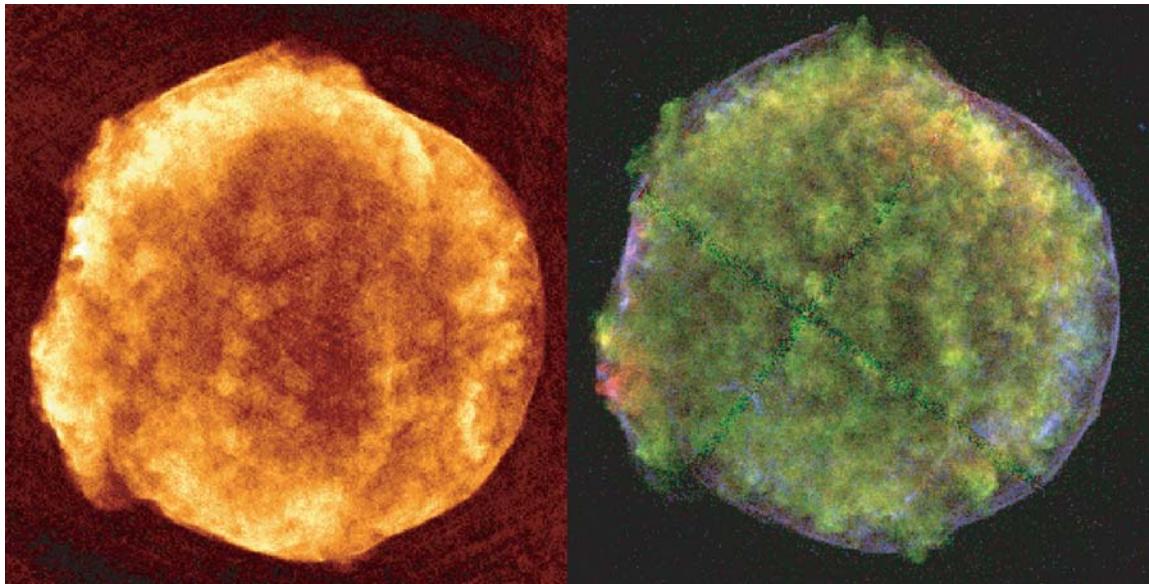


Plate 19 (left) A 1.375-GHz map of the remnant of Tycho's SN (diameter 8'). Colour signifies intensity, with white/yellow signifying bright, red faint. Beam size was 1.4''. Features correspond closely but not exactly with those in the X-ray map (Reynoso et al., 1997). (right) The remnant of Tycho's SN mapped by the Chandra ACIS detector. Colour here signifies photon energy. Red is 0.95–1.26 keV (from Fe); green is 1.63–2.26 keV (from Si); blue is 4.1–6.1 keV (continuum). Exposure was 145 ks. Note the sharp forward shock at the west (right) limb in both radio and X-ray bands (from Warren et al., 2005).

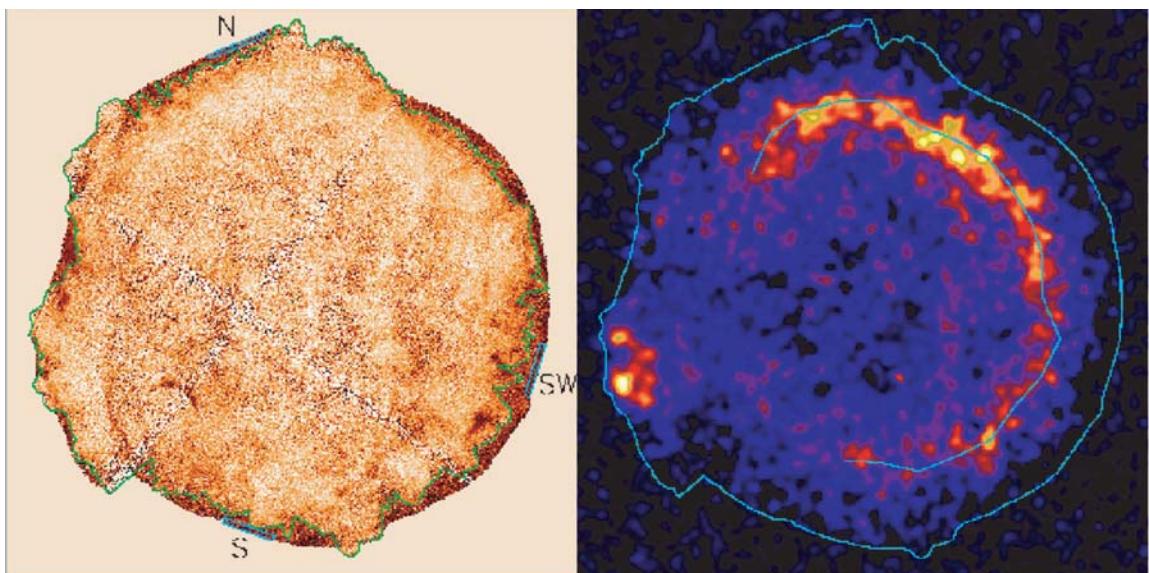


Plate 20 (left) X-rays from debris (pink and white) and from forward shock (dark red). The green contour is the contact discontinuity. (right) Material radiating from the Fe K line at 6–7 keV with continuum subtracted. Yellow is bright; blue is faint. Inner contour indicates the reverse shock. Outer contour shows the forward shock (from Warren et al., 2005).

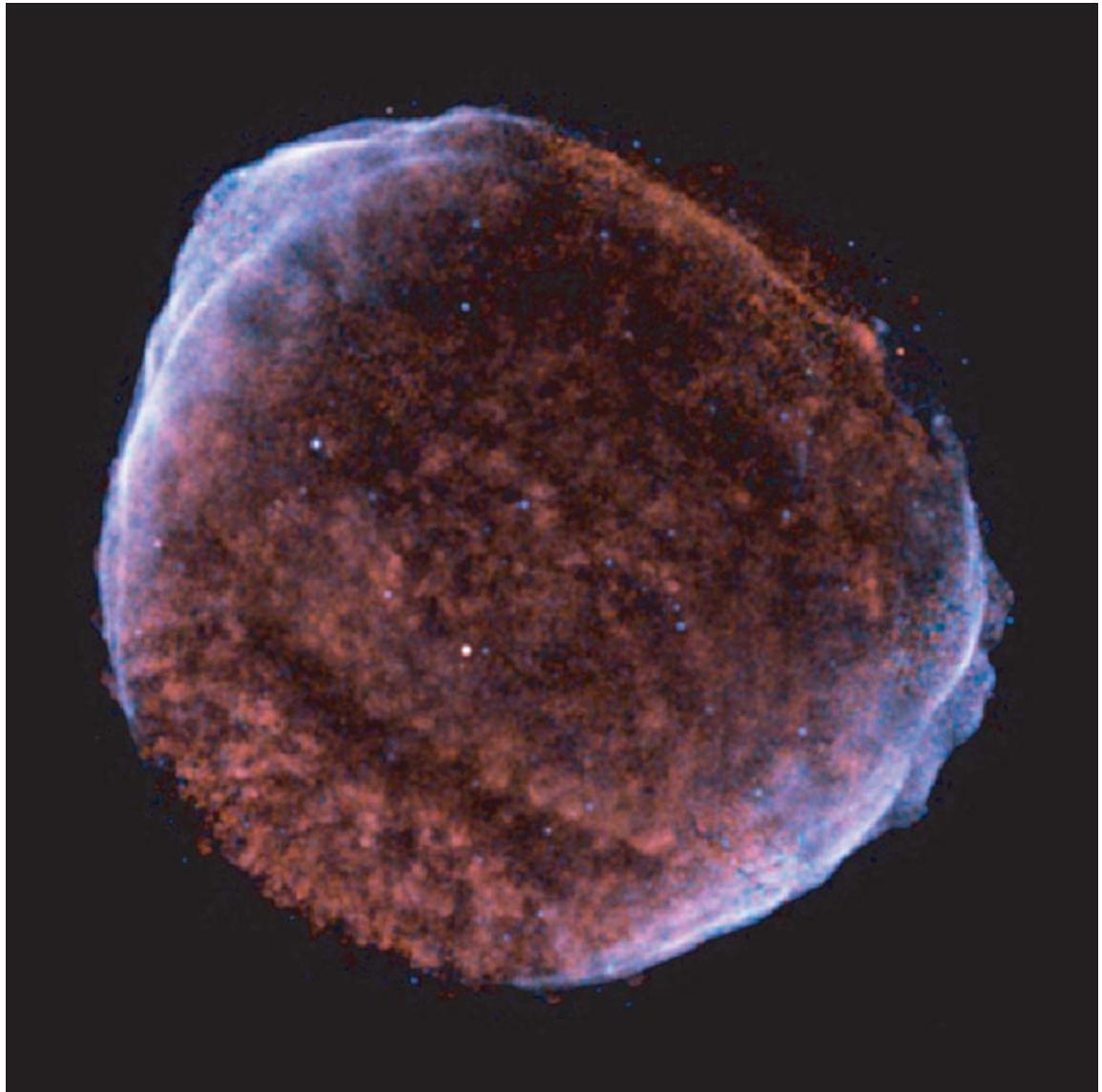


Plate 21 Chandra image of the remnant of SN 1006. The remnant is 36' in diameter. This is a mosaic of 11 ACIS exposures. Total observing time was 220 ks. Red shows photons in the range 0.5–0.9 keV. Cyan shows the energy range 0.9–1.3 keV. Blue shows the energy range 1.3–3.0 keV. Spectrum of X-rays from the blue regions is a continuum from high-energy electrons accelerated in the forward shock. Interior emission is from ejecta (NASA/CXC/Rutgers 2005; Cassam-Chenai et al., 2008; Long et al., 2003).

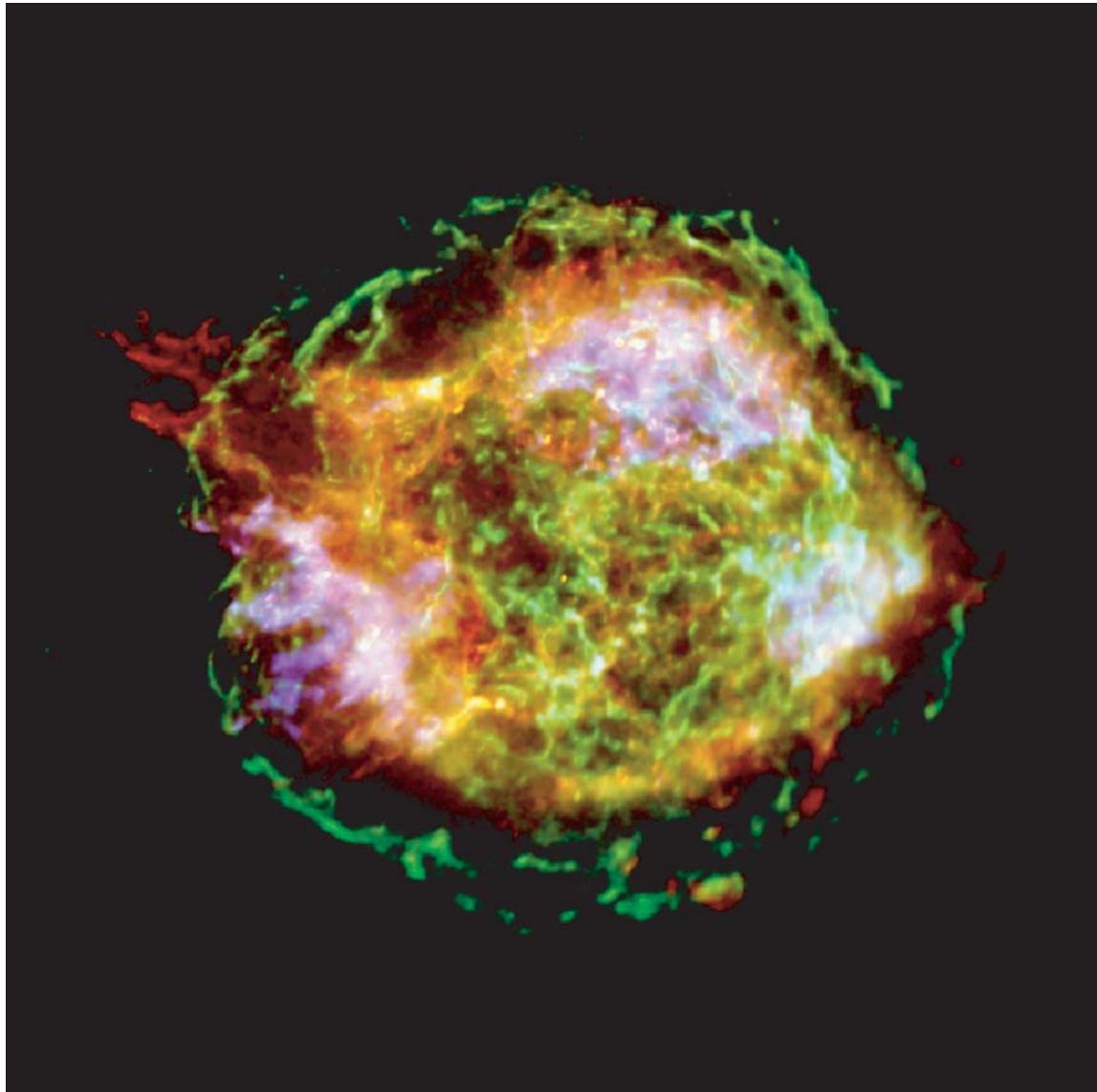


Plate 22 The Chandra 1 million second exposure of Cas A. Colour signifies photon energy. Red is from Si XIII, 1.8–2.0 keV. Blue is from Fe K, 6.5–7.0 keV. Green is continuum radiation, 4.2–6.4 keV. The ragged circular forward shock has a diameter of 4'. The ejecta are clumpy, and there is a prominent NE (upper left) jet which extends beyond the forward shock. Note the faint white point source at the centre (NASA/CXC/GSFC 2004; Hwang et al., 2004).

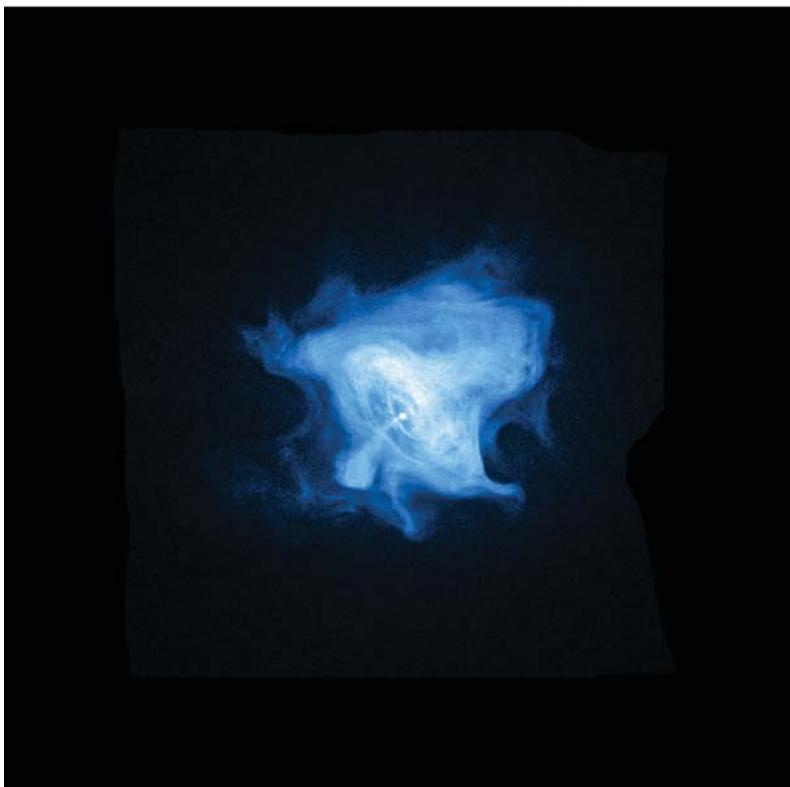
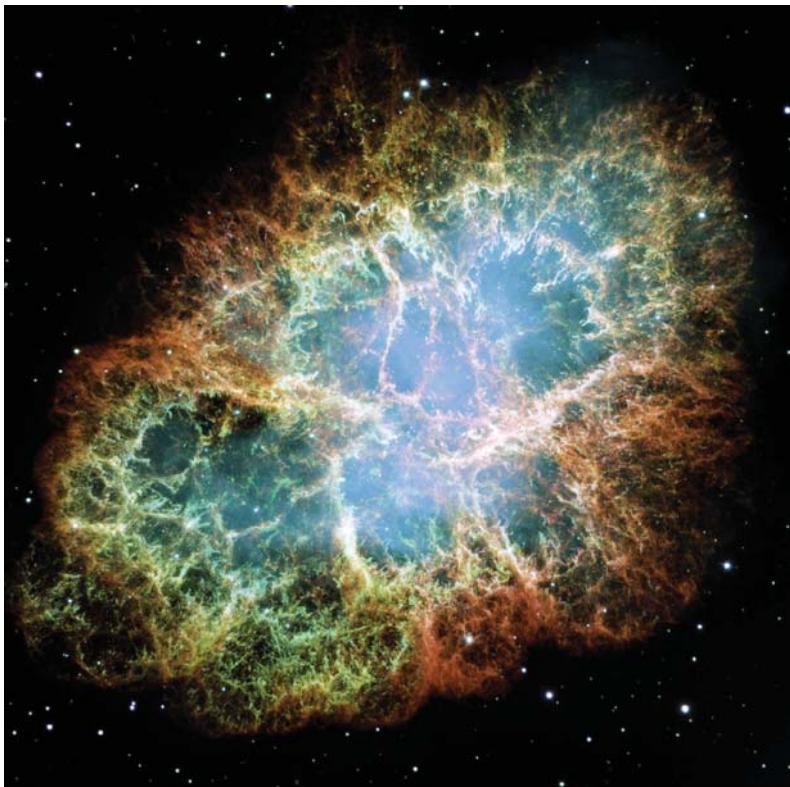


Plate 23 High-resolution pictures of the Crab Nebula at optical and X-ray wavelengths. Both panels are to the same scale – about 7' square. (top) A mosaic of 24 HST WFPC-2 images. The orange filaments show emission from material ejected from the star and the blue interior glow is a synchrotron continuum emitted by electrons accelerated by the central pulsar. (bottom) A Chandra ACIS image of the pulsar and surrounding PWN (NASA/ESA and Allison Loll/Jeff Hester, 2005, NASA/CXC/Seward et al., 2006).

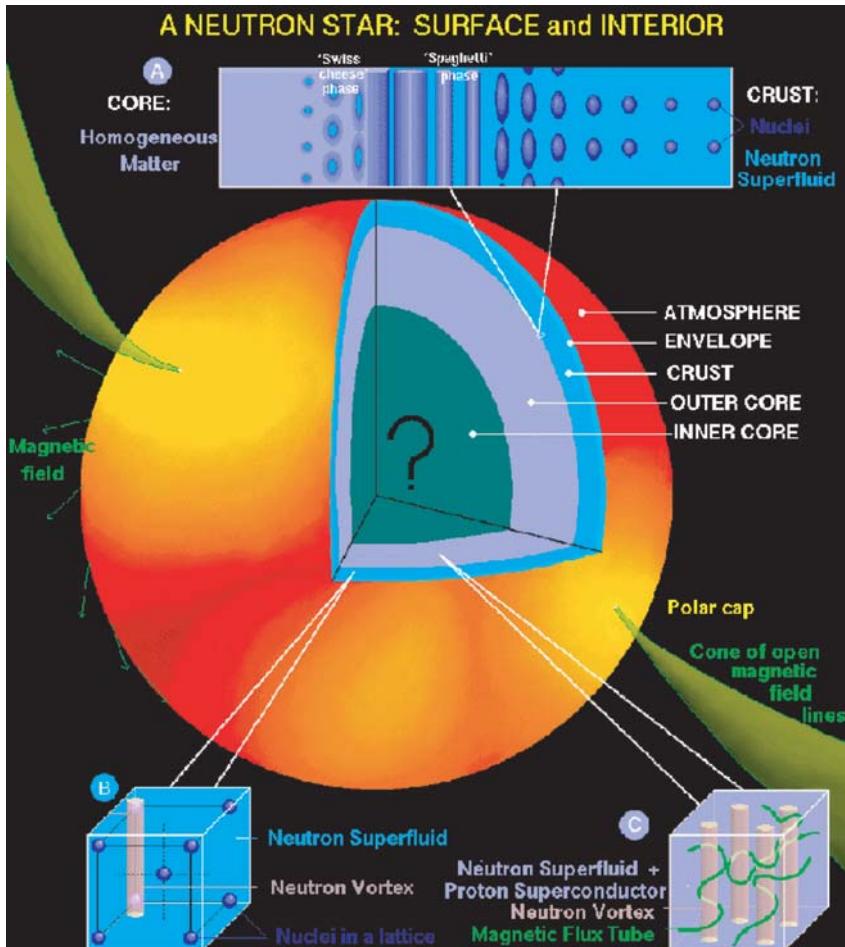


Plate 24 The structure of a neutron star. Insets show different configurations of nucleons and free neutrons. The magnetic field shown is not symmetrical and threads through the interior in vortex tubes which carry the angular momentum in the superfluid (courtesy Dany Page, 2009).

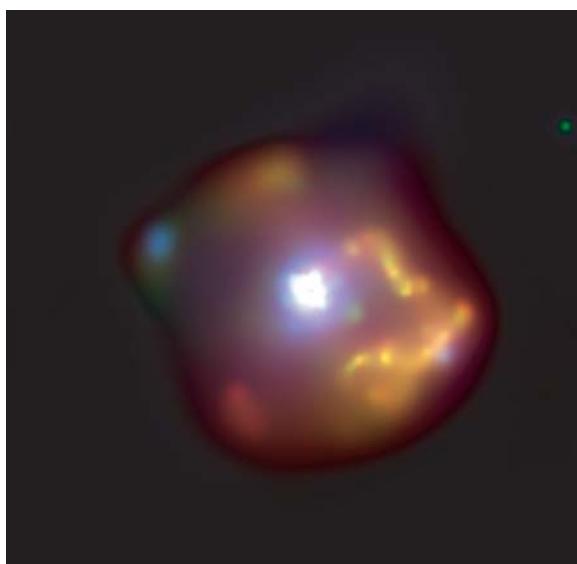


Plate 25 An 8-hour Chandra ACIS observation of SNR 0540-69 in the LMC. Colour indicates X-ray energy: red, 0.3–0.8; green, 0.8–2.1; blue, 2.1–4.0 keV. The uneven shell has a diameter of ≈ 1 arcmin and the emission is thermal. The bright PWN at the centre radiates more high-energy photons and surrounds a pulsar which is overexposed in this figure. The PSR/PWN combination is similar to that of the Crab Nebula (Seward et al., 2006).

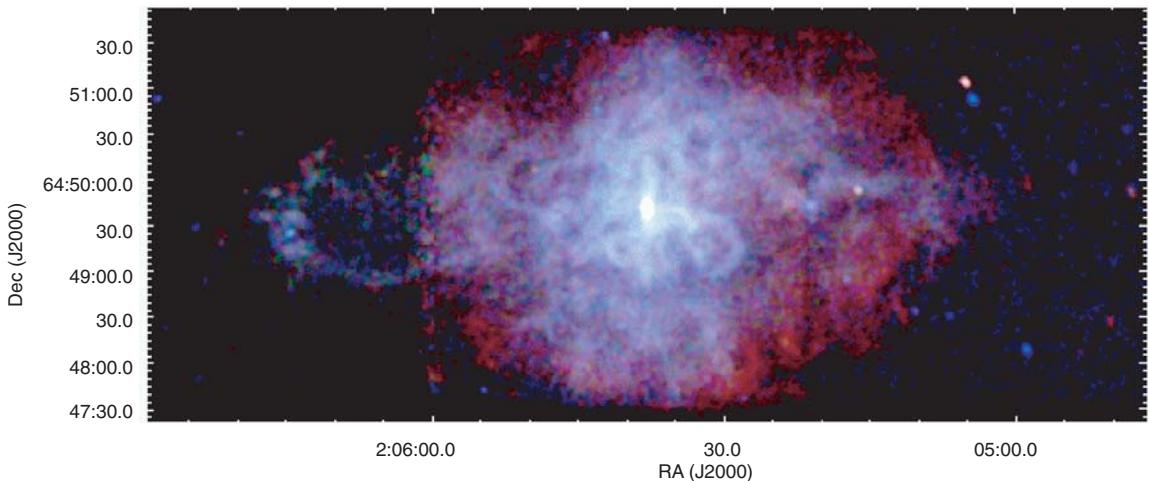


Plate 26 A 350-ks Chandra ACIS observation of the supernova remnant 3C 58. The image is 12.1×4.2 arcminutes in size. Colour indicates photon energy range: red is 0.5–1.0 keV; green is 1.0–1.5 keV; and blue is 1.5–10 keV. Thermal emission from the faint shell is soft and shows red. The more energetic PWN radiation is mostly blue, has a ropey appearance and fills the inside of the remnant. The bright centre of the PWN is a torus-oriented NS, which obscures the pulsar in this image. A possible jet extends westward along the axis of the torus (from NASA/CXC/Slane *et al.*, 2004).

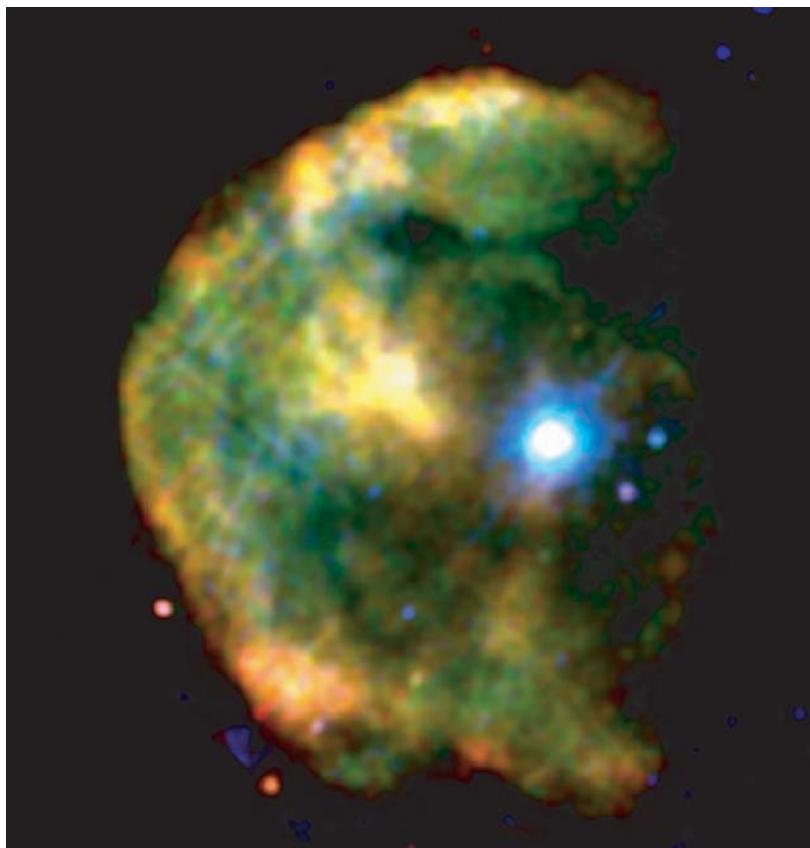


Plate 27 The supernova remnant CTB 109 as seen by the XMM EPIC detector. This is a mosaic of four exposures of duration 3–5 hours. Colour is used to identify X-ray energy: red is 0.3–0.9 keV; green is 0.9–1.5 keV; and blue is 1.5–4.0 keV. The outer shock forms a remarkable half-circle over the east half of the remnant. There is a molecular cloud to the west where the shock has evolved to invisibility. Colour shows the bright AXp at the centre of the remnant to be harder than the other emission. Note that there is no PWN around the AXp which, in this image, because of overexposure and smoothing, appears larger than its true point-source size (from Sasaki *et al.*, 2004).

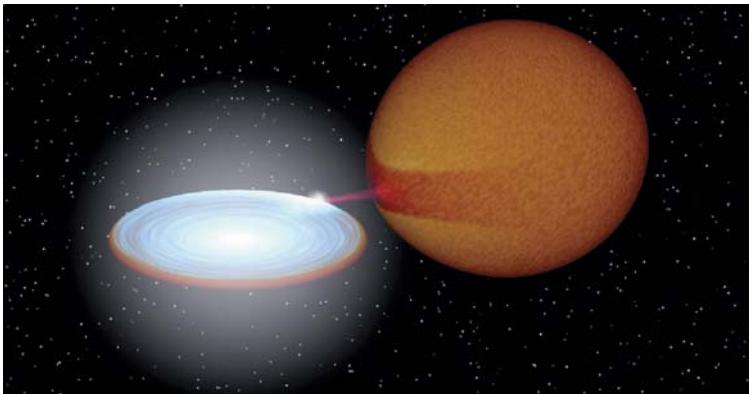


Plate 28 Visualisation of a CV, involving realistic representations of the structure and relative brightness of the different components. Note the mass-transfer stream from the inner Lagrangian point of the larger but cool secondary star, which gives rise to a ‘bright spot’ on the edge of the accretion disc. Since the bulk of the luminosity comes from the hot, inner accretion disc, the outer disc throws a ‘shadow’ onto the equatorial regions of the companion (figure by Rob Hynes).

Plate 29 The X-ray emission line components of Vela X-1 from Fig. 11.25 showing that the line fluxes are essentially constant as a function of orbital phase. Note that the $\lambda 6.5\text{--}7 \text{\AA}$ region of the spectrum (all due to Si II – Si XIII ions) appeared as essentially a single, slightly broadened feature in the first ASCA spectra of this source (diagram by Sako *et al.*, 2003).

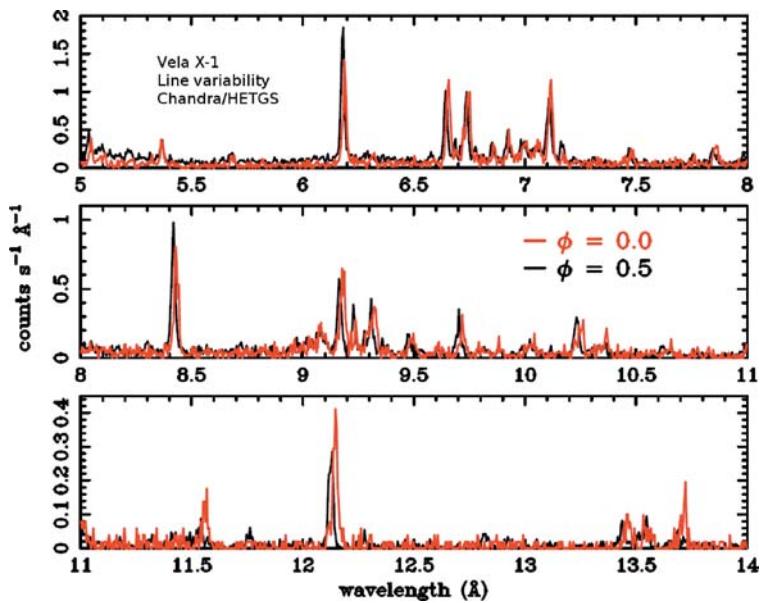


Plate 30 Artist’s impression of the SS433 binary system, showing all the important components (painting by Bill Hartmann).



M80
(HST)

Plate 31 Colour photograph with HST's WFPC2 camera of the globular cluster M80, an eighth-magnitude object visible through binoculars in the constellation of Scorpius and at a distance of about 10 kpc. Note how the density of stars increases rapidly towards the centre of the cluster, there being several hundred thousand in all. All these stars are less massive than our Sun, but the brightest have evolved to become red giants. However, the low metal abundance in globular clusters (due to their age) means that the stars will be much bluer than their galactic counterparts (image from the Hubble Heritage Team, AURA/STScI/NASA).

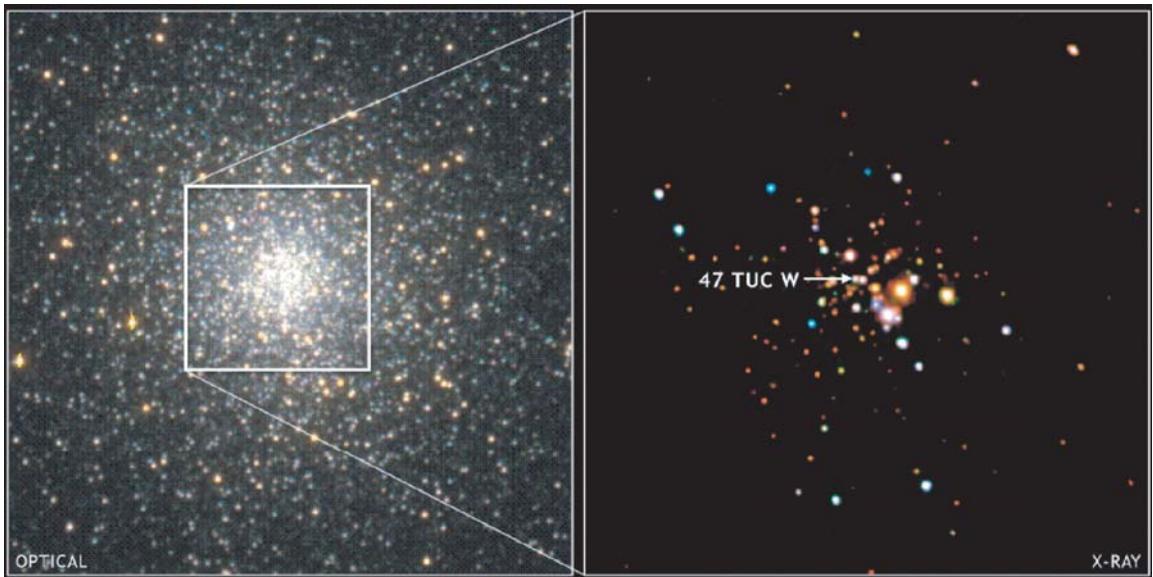


Plate 32 (left) The naked-eye globular cluster 47 Tucanae is, at 5.1 kpc, half the distance away of M15 and strongly centrally condensed. This ground-based image (with ESO's 1.54 m telescope) covers 6.6 arcmin square, and the central 2.5-arcmin square is imaged by Chandra's ACIS (right). The colours reflect the X-ray energies, with 0.3–1.2 keV red, 1.2–2 keV green and 2–6 keV blue. The entire gamut of low- L_X sources (Fig 11.84) are present here, with the brightest MSP marked out of the 19 MSPs detected in the cluster by Chandra (X-ray: NASA/CXC/CfA/J. Grindlay & C. Heinke; optical: ESO/Danish 1.54-m/W. Keel et al.).

Black Hole Binaries in the Milky Way

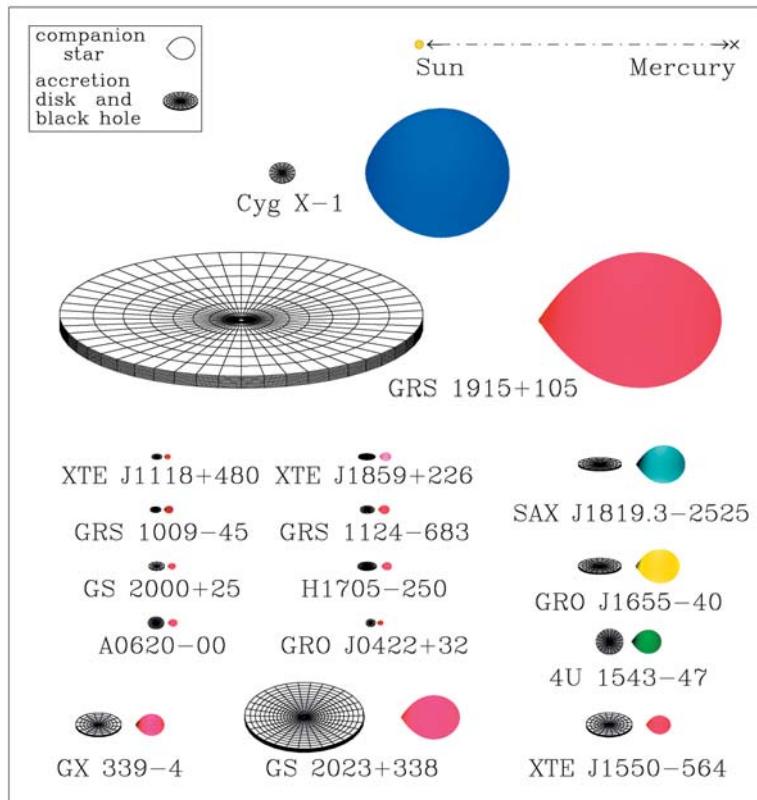


Plate 33 Binary schematics (to scale) of 16 of the galactic black-hole X-ray binaries, with the distance from the Sun to Mercury shown at the top for scale. The accretion discs are tilted so as to show the measured orbital inclinations, and the donor star colour gives an indication of its surface temperature (compiled by Jerry Orosz).

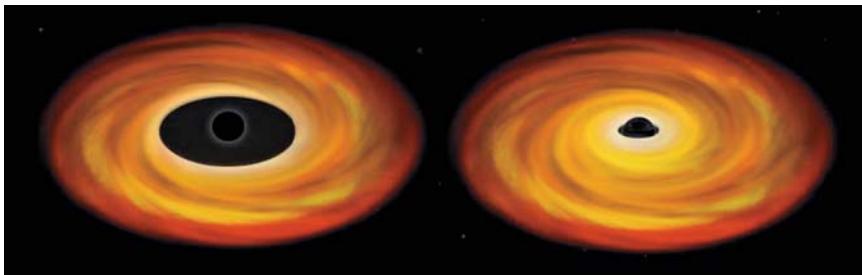


Plate 34 Schematic of the accretion disc surrounding a black hole, with a central ‘empty’ region defined by the radius of the innermost stable circular orbit, within which material must spiral rapidly into the black hole. (left) If the black hole is not spinning, then it is a Schwarzschild BH, and the innermost material is cool. (right) However, a maximally spinning BH is a Kerr BH with a much smaller innermost radius, and the material can be much hotter.

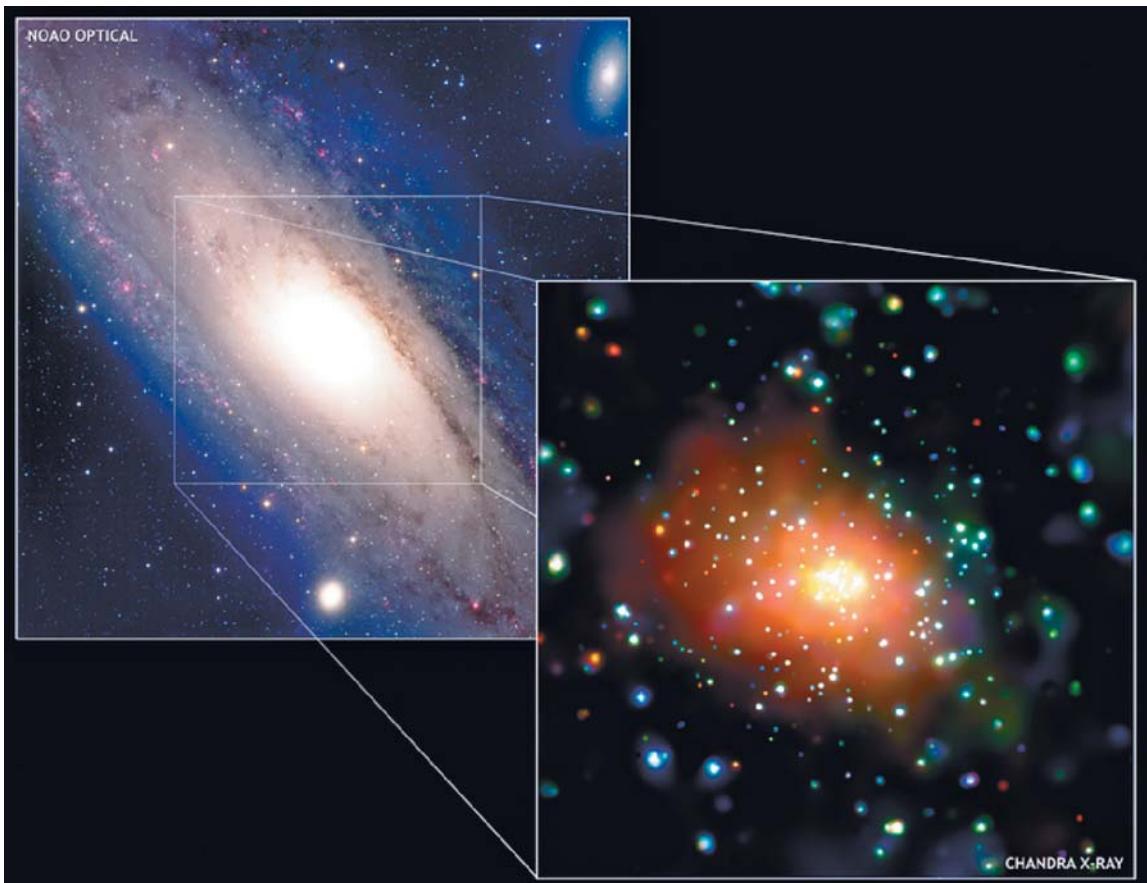


Plate 35 An optical picture of M31 compared with the central region imaged by Chandra. The Chandra field is 28' square. Many pointings totalling 54 hours have been merged for this picture. Energy is colour coded: red, 0.5–1.0 keV; green, 1.0–2.0 keV; blue, 2.0–4.0 keV (X-ray: NASA/UMass/Z. Li and Q. D. Wang; Optical: NOAO/AURA/NSF/T. Rector and B. Wolpa; NASA/CXC 2006).

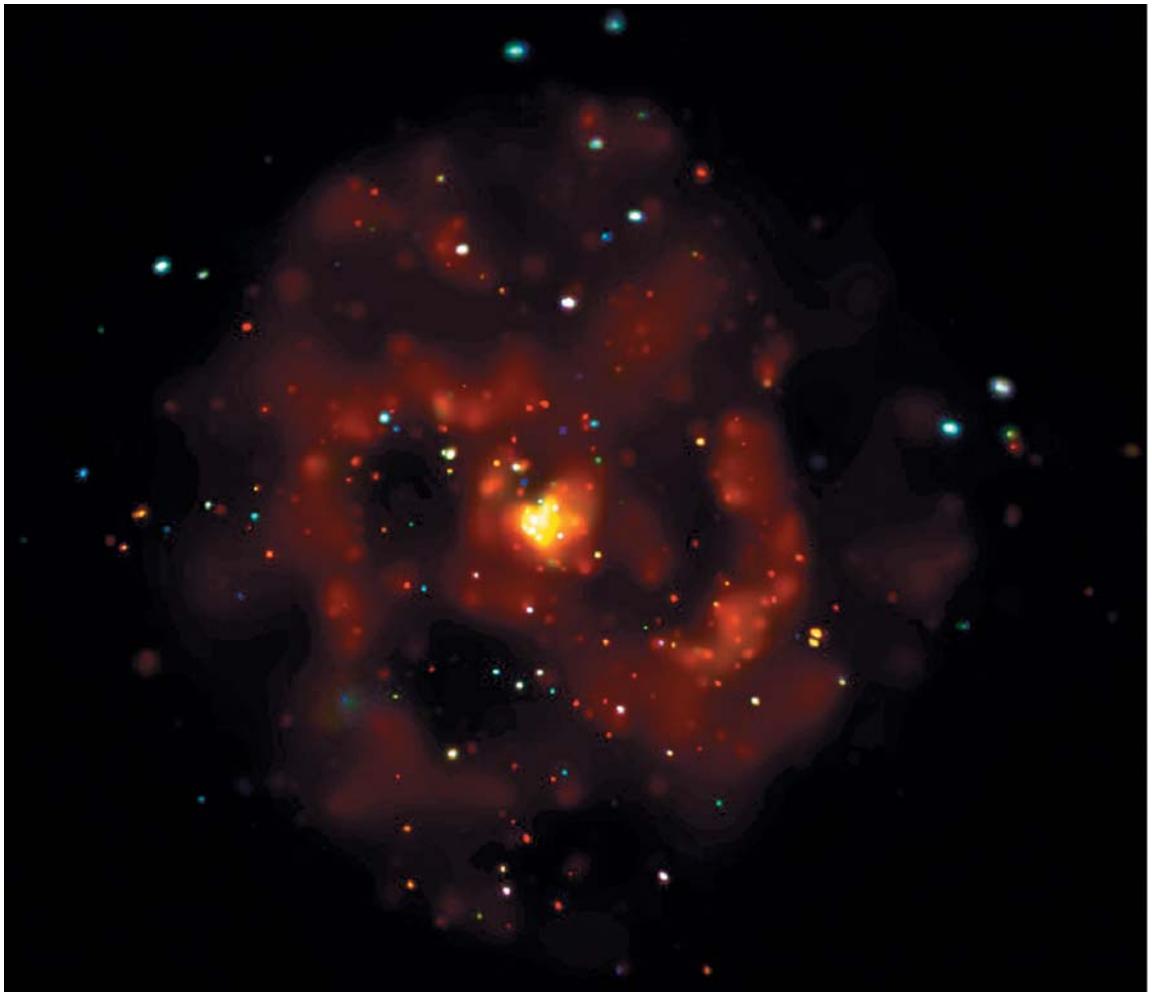


Plate 36 The spiral galaxy M83 imaged by Chandra. Colour indicates energy: red, 0.3–1 keV; green, 1–2 keV; blue, is 2–8 keV (NASA/CXC/U. Leicester/U. London, Soria & Wu, 2003; NASA/CXC 2003).

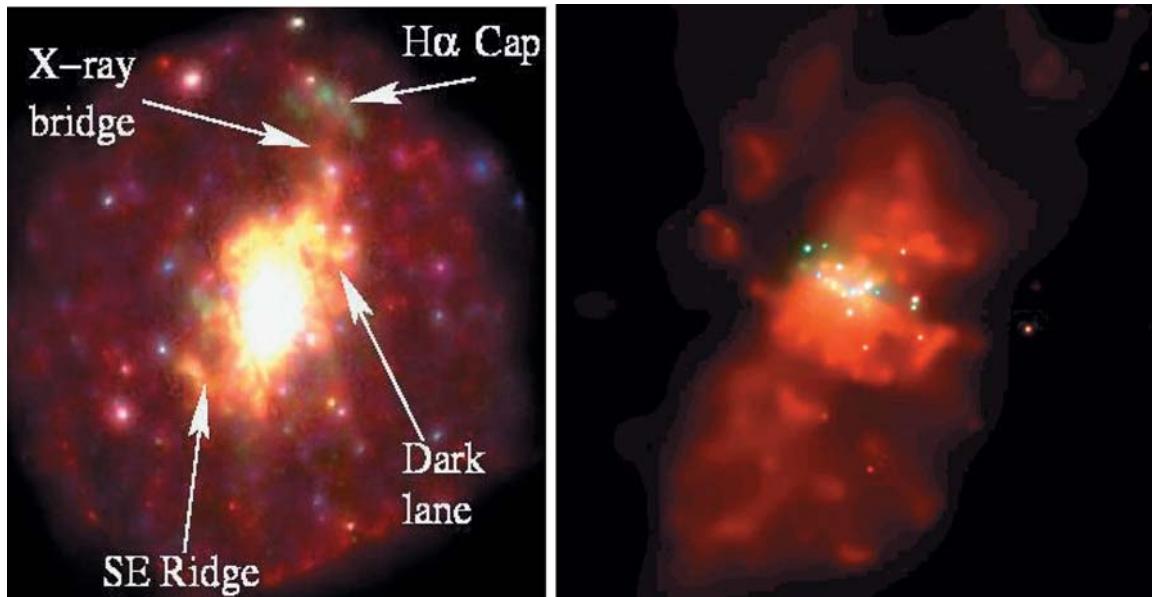


Plate 37 M82 imaged by XMM and by Chandra. (left) The XMM field is 26' square. The smaller Chandra field, (right) is 5' square and covers only the bright centre of the galaxy. Exposures are 31 and 35 ks, respectively. Energy in the XMM image is colour coded: red, 0.2–0.5; green, 0.5–0.9; blue, 0.9–2.0 keV. Colour coding in the Chandra data is similar (Stevens *et al.*, 2004; NASA/SAO/Fabbiano *et al.*, 2001).



Plate 38 A mosaic of Chandra fields showing a $0.8^\circ \times 2.0^\circ$ strip of the galactic plane centred on SgrA. Colour indicates energy: red, 1–3 keV; green, 3–5 keV; blue, 5–8 keV. Two of the bright sources in Fig. 13.11 appear large and overexposed in the plane, and the halo of a third, A1742-294, appears at the bottom (NASA/CXC/U. Mass; Wang *et al.*, 2002; NASA/CXC 2002).

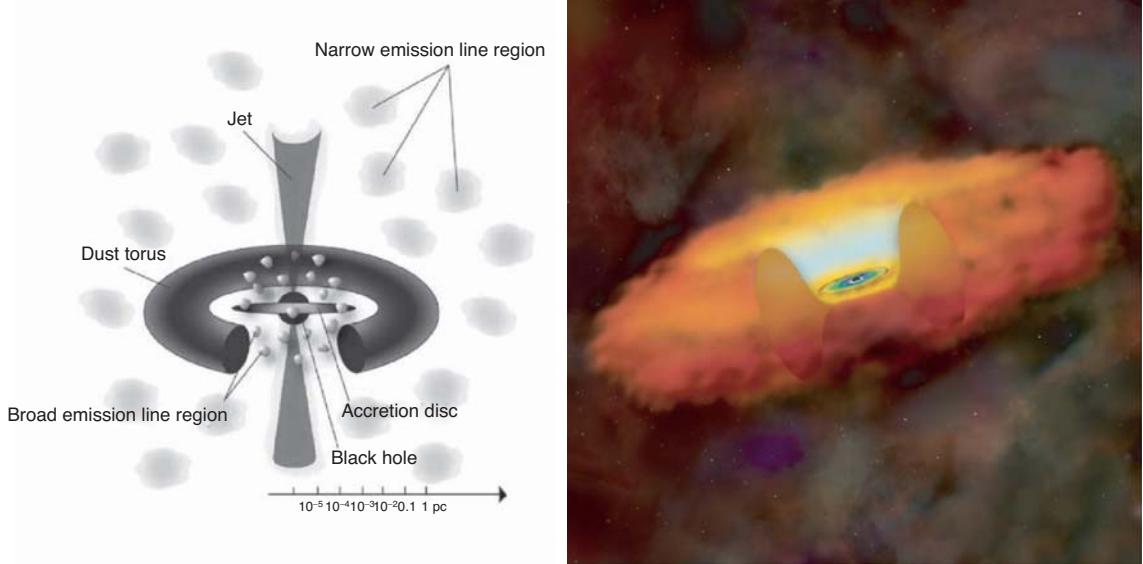


Plate 39 The classic 'unified' model of an AGN (right) as depicted in cutaway form and (left) with individual components labelled.



Plate 40 Multi-wavelength image of NGC 4258 assembled from a combination of space (Chandra, in blue; Spitzer, in red) and ground-based observations (digital sky survey, in yellow) (image courtesy NASA, Harvard-Smithsonian CfA).

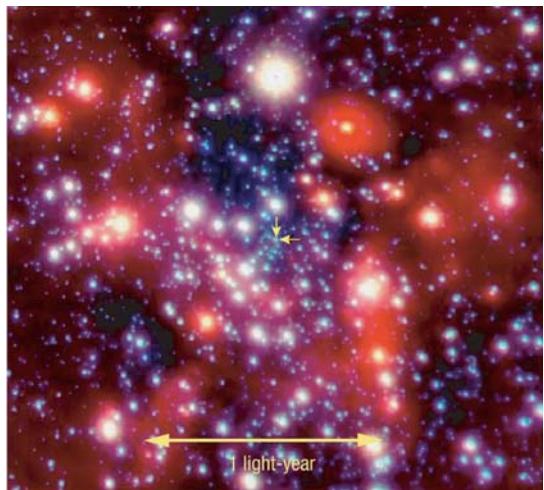


Plate 41 Adaptive optics-corrected IR image of the Galactic Centre obtained with the ESO VLT. The position of the radio source, Sgr A*, defines the Galactic Centre and is arrowed at the centre of the image. The scale bar of 1 light year demonstrates how small a volume of space is covered by this image, yet more than 1000 stars can be identified within that distance of the Galactic Centre. A combination of images through different IR filters between 1.6 and 3.5μ produces the colours, with blue indicating hotter and red cooler stars (image courtesy of ESO).

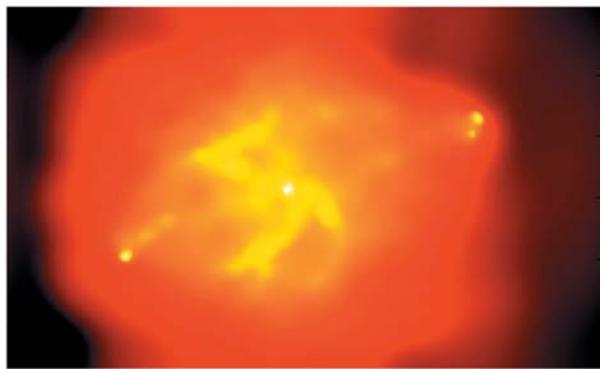
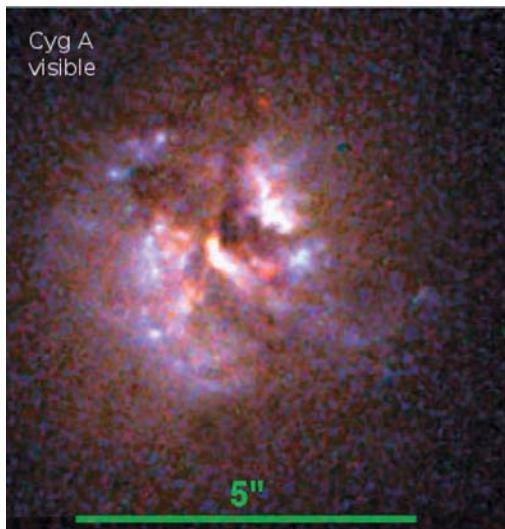


Plate 42 (left) Combined Keck and HST visible image of Cyg A, showing complex absorption/scattering structure around the elliptical galaxy. (right) Chandra X-ray image shows that the hot spots (also seen in the radio maps) are where the jets impact the surrounding hot intergalactic medium of the cluster in which Cyg A resides (images courtesy of ESO and NASA/UMD/A. Wilson et al.).

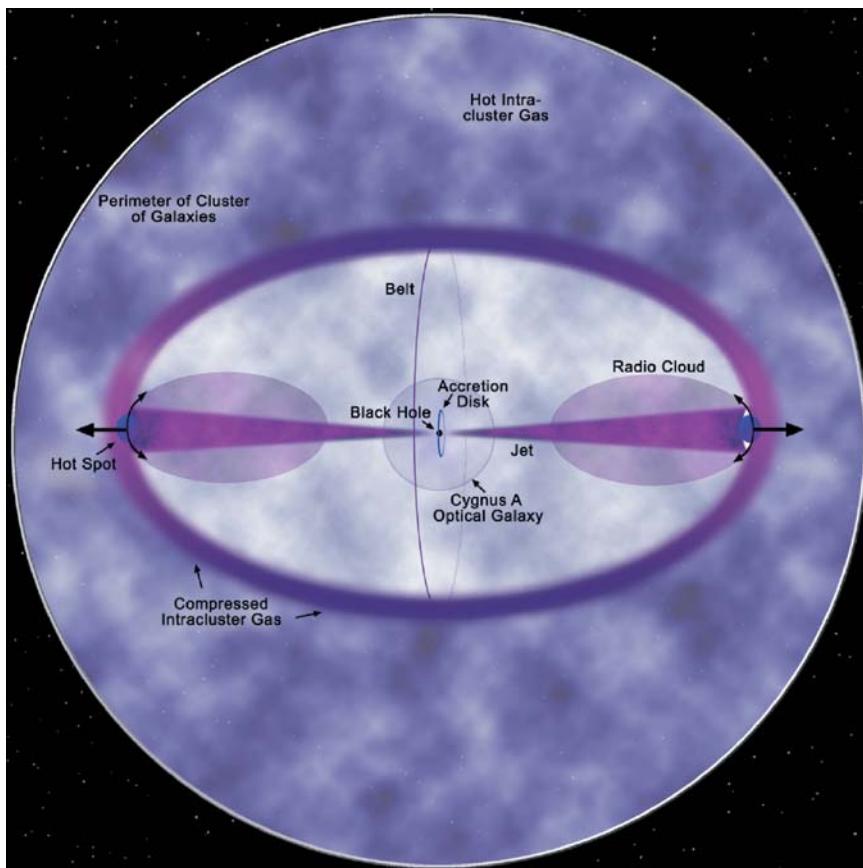


Plate 43 Illustration of the various observed radio, optical and X-ray components of Cyg A. Note in particular how the jets have evacuated the inner region, producing the X-ray ‘cavity’ (image courtesy NASA/CXC/K. Kowal).

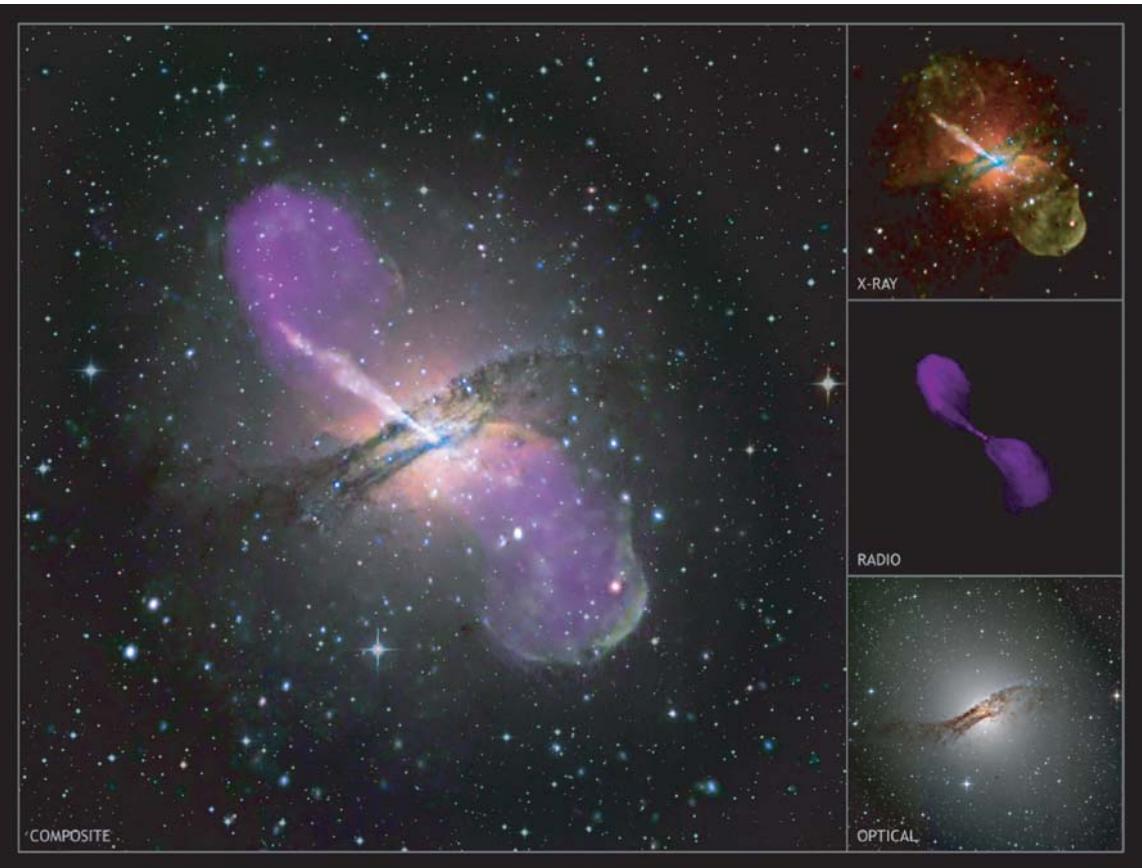


Plate 44 Multi-wavelength composite image of the nearest AGN, Cen A, with the individual (top) Chandra, (middle) radio and (bottom) optical images shown on the right (X-ray: NASA/CXC/CFA/R. Kraft *et al.*; Radio: NSF/VLA/UvA Hertfordshire/M. Hardcastle; Optical: ESO/WFI/M. Rajtuba *et al.*).

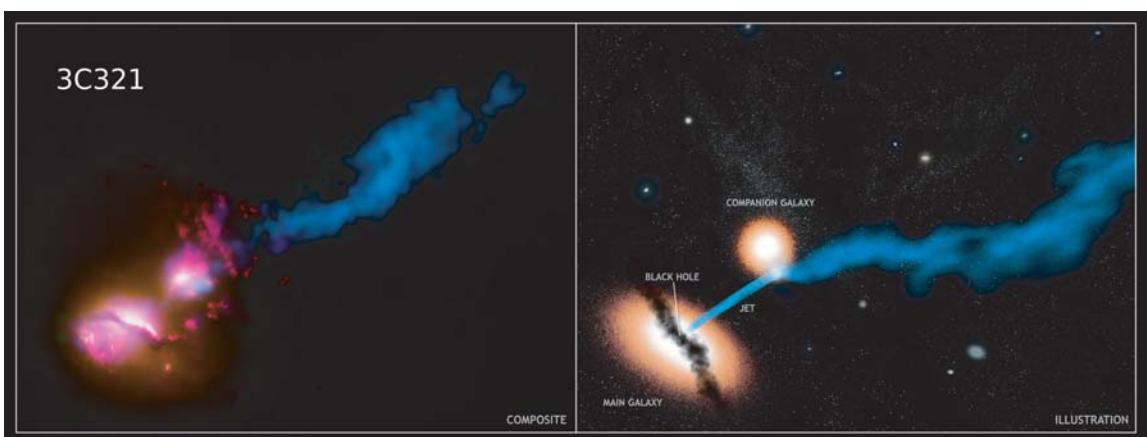


Plate 45 (left) Composite X-ray (Chandra, purple), optical (HST, yellow) and radio (VLA/MERLIN, blue) image of the radio galaxy 3C321. The schematic (right) shows how the radio jet from the SMBH in the radio galaxy 3C321 is impacting a smaller, neighbouring galaxy (only about 6 kpc away), producing a local hot spot and some deflection of the jet (X-ray: NASA/CXC/CFA/O. Evans *et al.*; Optical/JV: NASA/STScI; Radio: NSF/VLA/CFA/O. Evans *et al.*, STFC/JBO/MERLIN).



Plate 46 Combined X-ray and radio image of 3C75 in the Abell 400 cluster of galaxies. The radio image is of the jets from a pair of SMBHs, and these are embedded in the hot (10^7 K), diffuse X-ray-emitting gas of the cluster. The pair of SMBHs are only ~ 7 kpc apart, and the galaxy in which they are embedded (NGC 1128) has a dumb-bell appearance, suggesting a merging galaxy pair (X-ray: NASA/CXC/Alfa/D. Hudson & T. Reiprich et al.; Radio: NRAO/VLA/NRL).

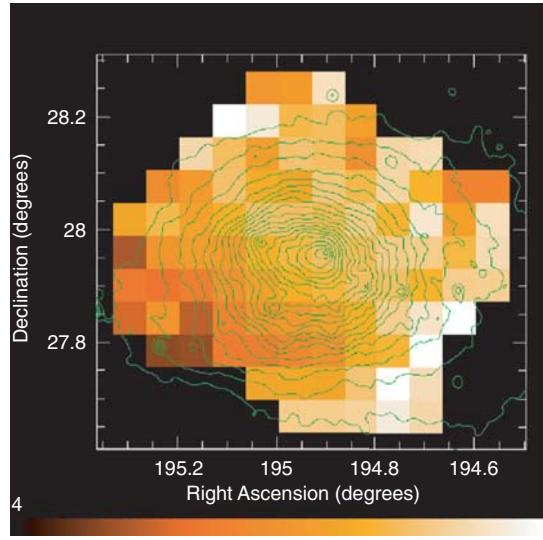


Plate 47 XMM EPIC-MOS temperature map of the Coma Cluster central region. The field is $42' \times 45'$. White is hot (11 keV) and brown/red is cool (6 keV). Boxes are $3.5'$ square. Isointensity contours from Colour Plate 48 are overlaid (from Arnaud et al., 2001).

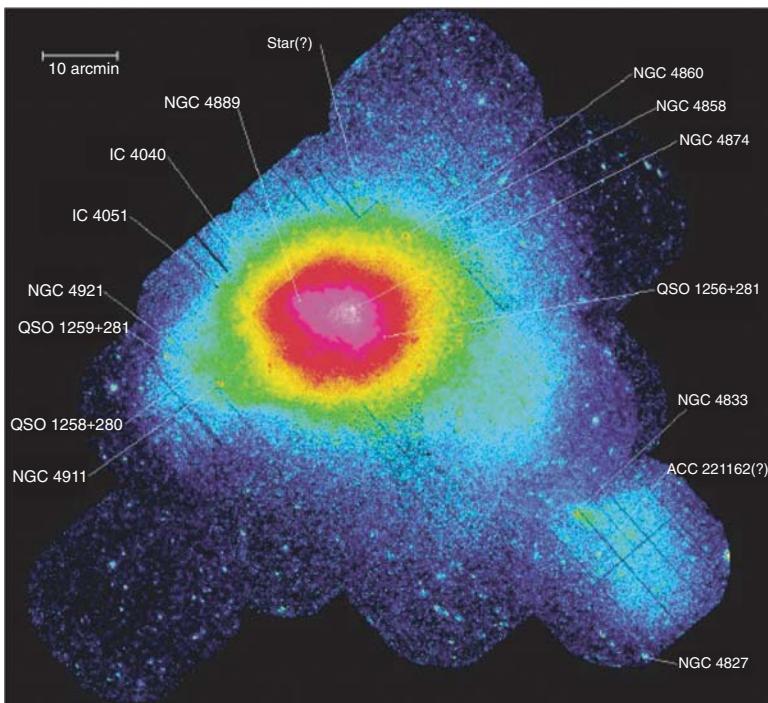


Plate 48 An XMM X-ray image (0.3–2 keV) of the Coma Cluster. Intensity is colour coded with purple/red being brightest. This is a mosaic of 12 EPIC-pn images. Exposures average ~ 20 ks. Brighter point sources are identified. The principal emission visible here is $\sim 40' \times 55'$ in extent (1.2×1.6 Mpc at the cluster). The large blob in the lower right is from the in-falling NGC 4839 group (from Briel et al., 2001).

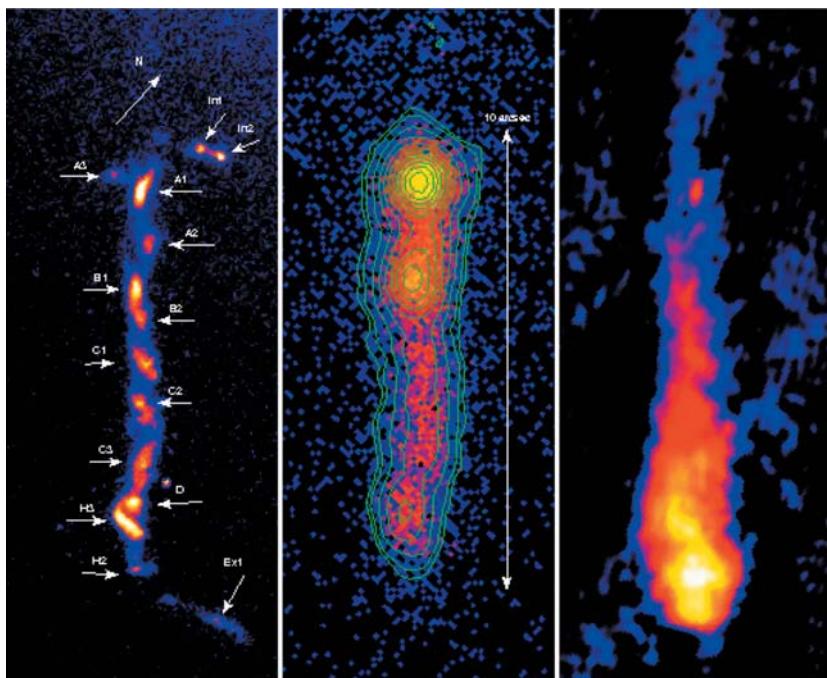


Plate 49 Blowup of the jet of 3C273 as seen at (left) optical (HST), (middle) X-ray (CXO) and (right) radio (MERLIN), all to the same scale. The particles are propagating downwards (image courtesy NASA/STScI [optical, radio], NASA/CXC [X-ray]).



Plate 50 A smoothed Chandra ACIS-I X-ray image of the Bullet Cluster (IE 0657-56). The field is $6' \times 8'$ (2.4×3.2 Mpc at the cluster). X-rays are in red, optical in yellow and mass distribution from a lensing analysis in blue. (X-ray: NASA/CXC/CfA/M. Markevitch et al.; optical and lensing: NASA/STScI; ESO VVFI; Magellan/U. Arizona/D. Clowe et al.; NASA/CXC 2006).

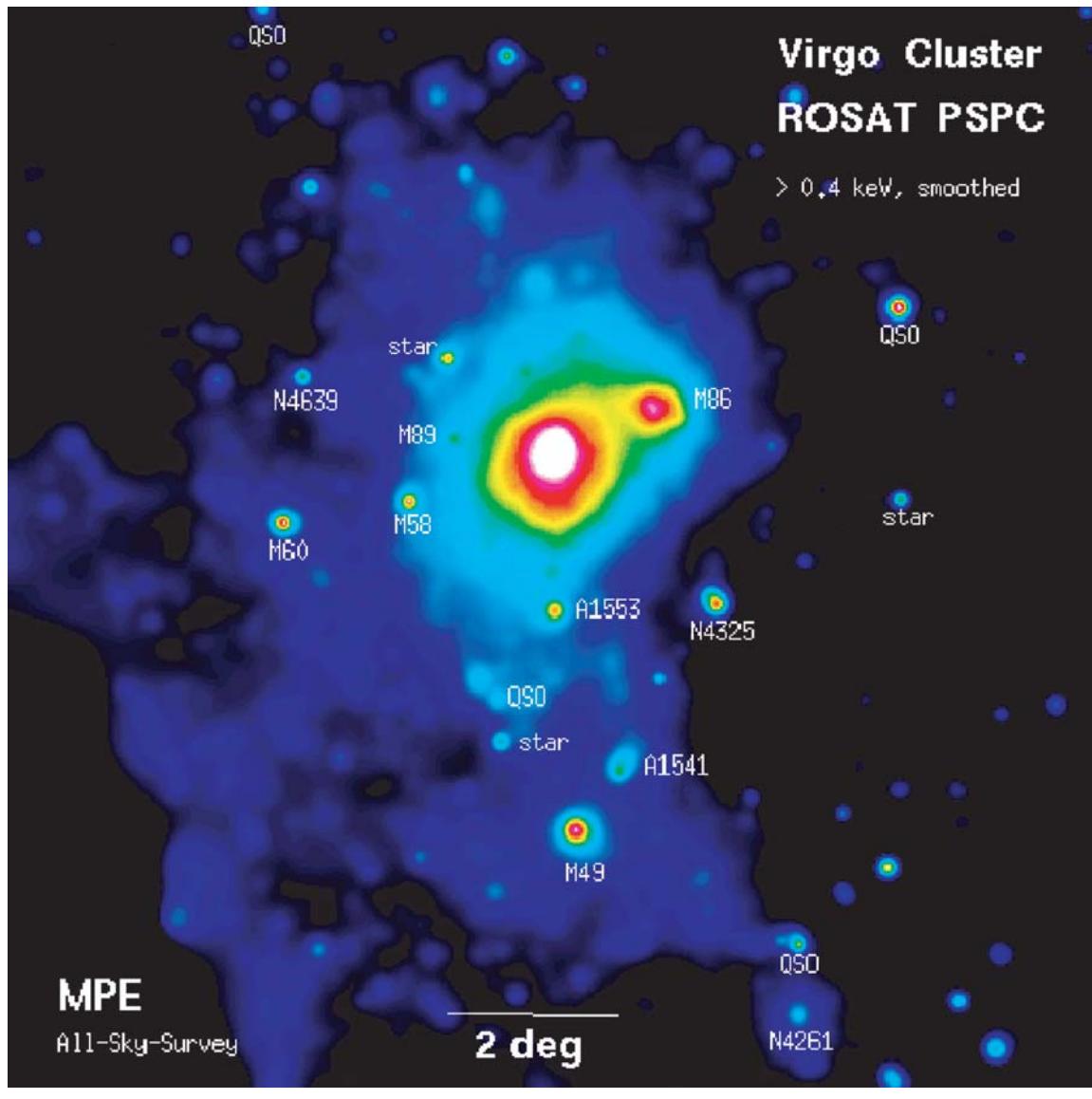


Plate 51 The ROSAT survey image of the Virgo Cluster. The field is 12° square (3.3 Mpc at the cluster). Colours indicate intensity. Individual labelled sources are galaxies in the cluster, foreground stars and background clusters and quasars. The bright central source, responsible for most of the emission, is the dominant galaxy M87. Emission 1.5° to the northwest is from an in-falling group containing the galaxy M86 (NGC 4406). Diffuse emission 4°–5° south of M87 is roughly centred on the galaxy M49 (NGC 4472) (Böhringer et al., 1994).

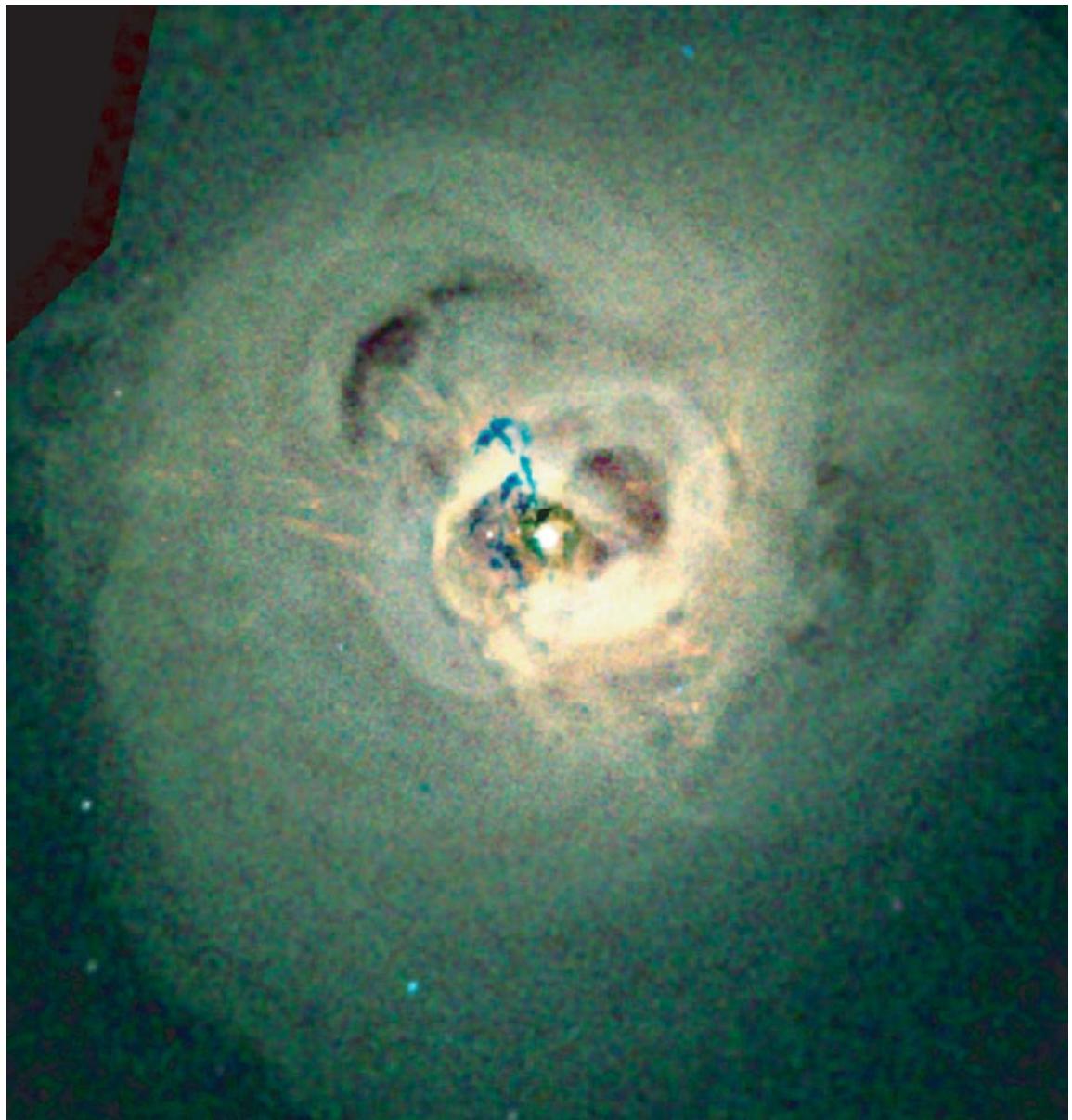


Plate 52 A one-million-second Chandra observation of the centre of the Perseus Cluster. Colour signifies energy: red is 0.3–1.2 keV, green is 1.2–2 keV and blue is 2–7 keV. The structure seen in Fig. 15.11 is now visible without special processing. The field is 6' square. This is 130 kpc at the cluster, a smaller region than for the other clusters shown. You are looking at the core of the cluster. The blue-green structure close to the centre is caused by absorption in in-falling material (NASA/CXC 2005; Fabian et al., 2006).

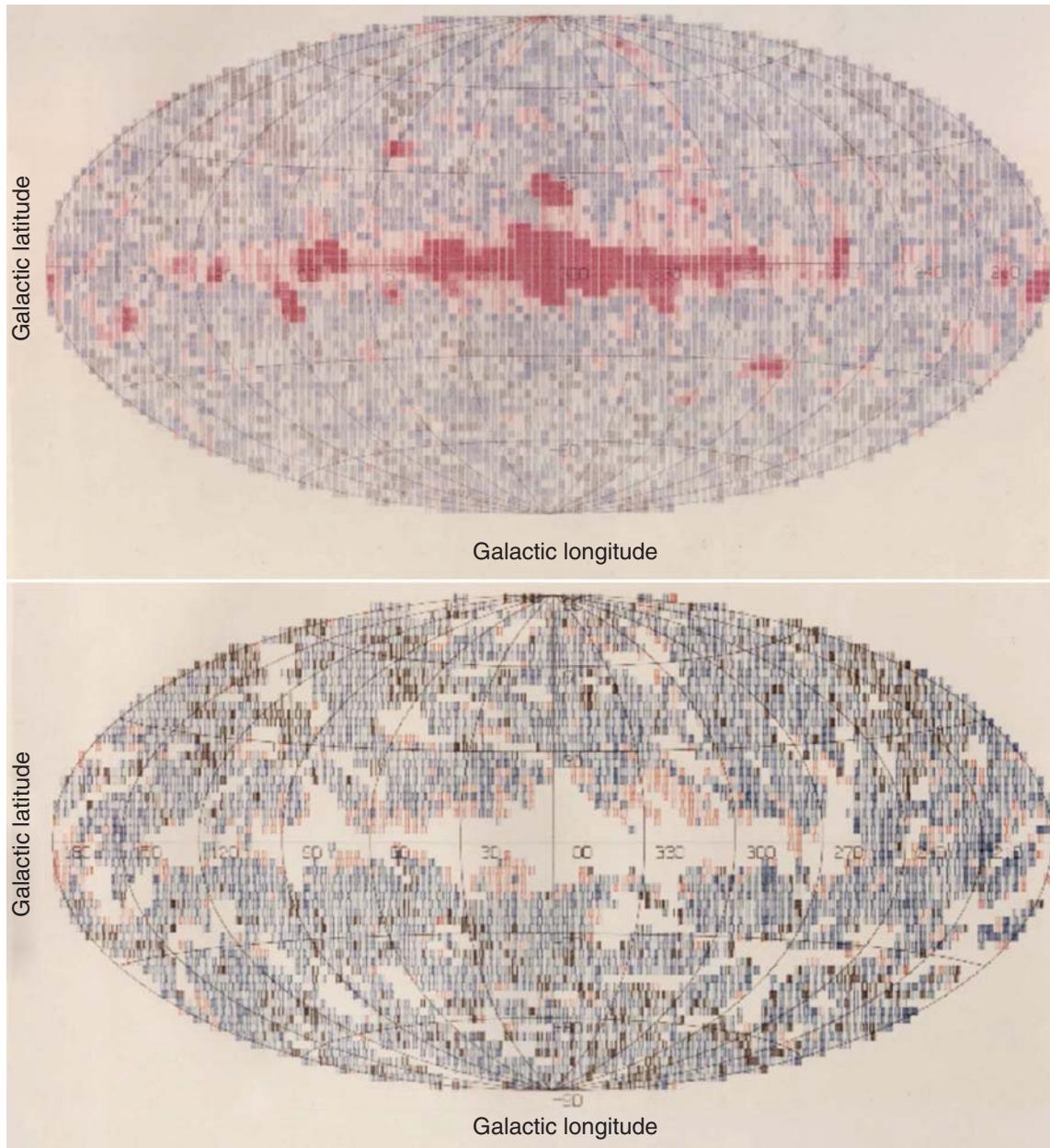


Plate 53 (top) The surface brightness of the X-ray sky mapped by HEAO-1. Pixels are $3^\circ \times 3^\circ$, energy range is 2–60 keV and the map is in galactic coordinates and centred on the Galactic Centre. The sky brightness is colour coded black (lowest), blue, pink and red (highest). The brightest emission is from discrete sources lying in the galactic plane and concentrated toward the Galactic Centre. (bottom) Resolved sources have been removed and pixels increased to $9^\circ \times 9^\circ$. Each line per pixel represents 1.5 per cent of the mean intensity, which is at about eight lines on this map. The dark pixels show unresolved point sources. Note that the emission is brightest close to the galactic plane (courtesy F. Marshall, GSFC).

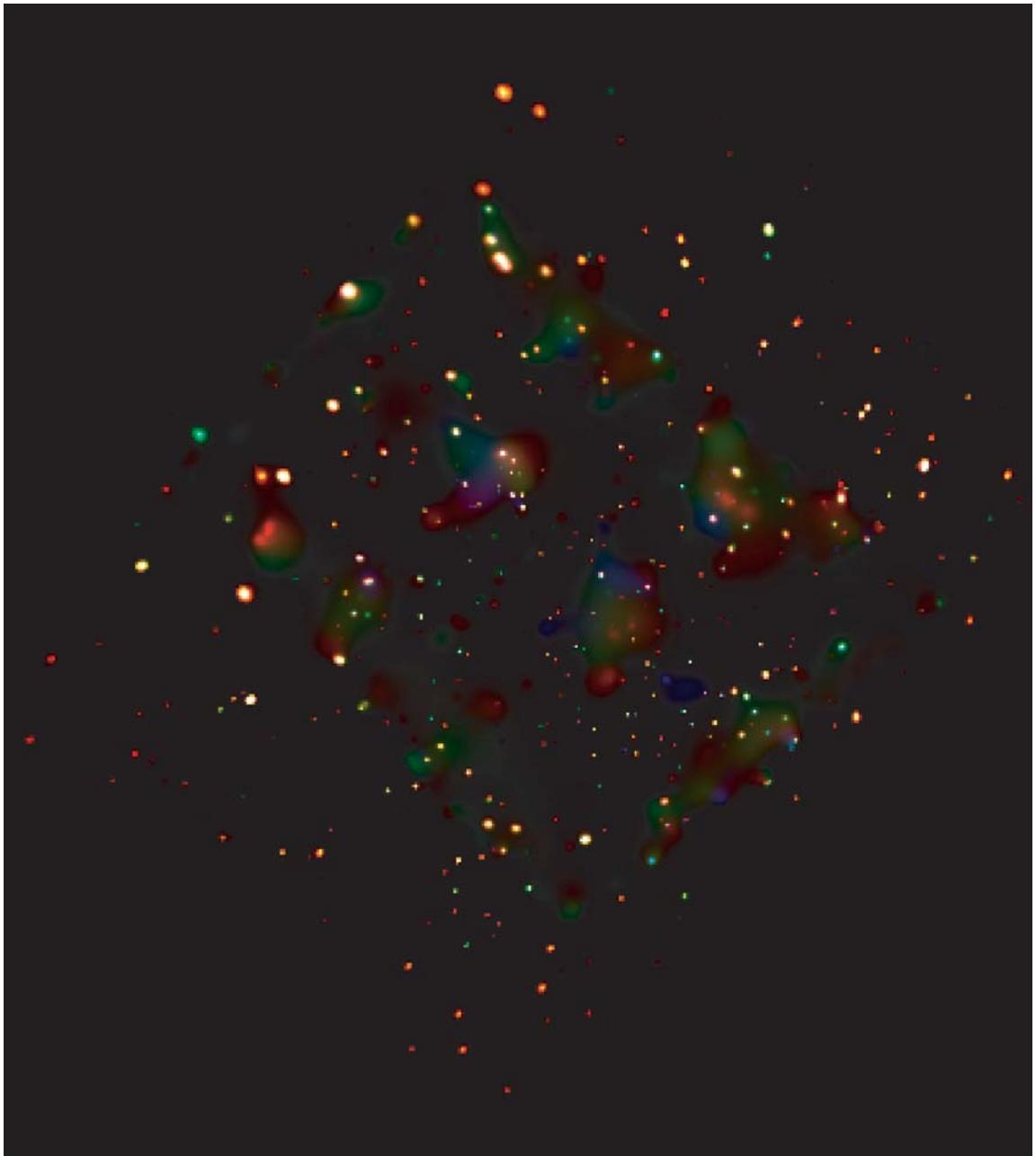
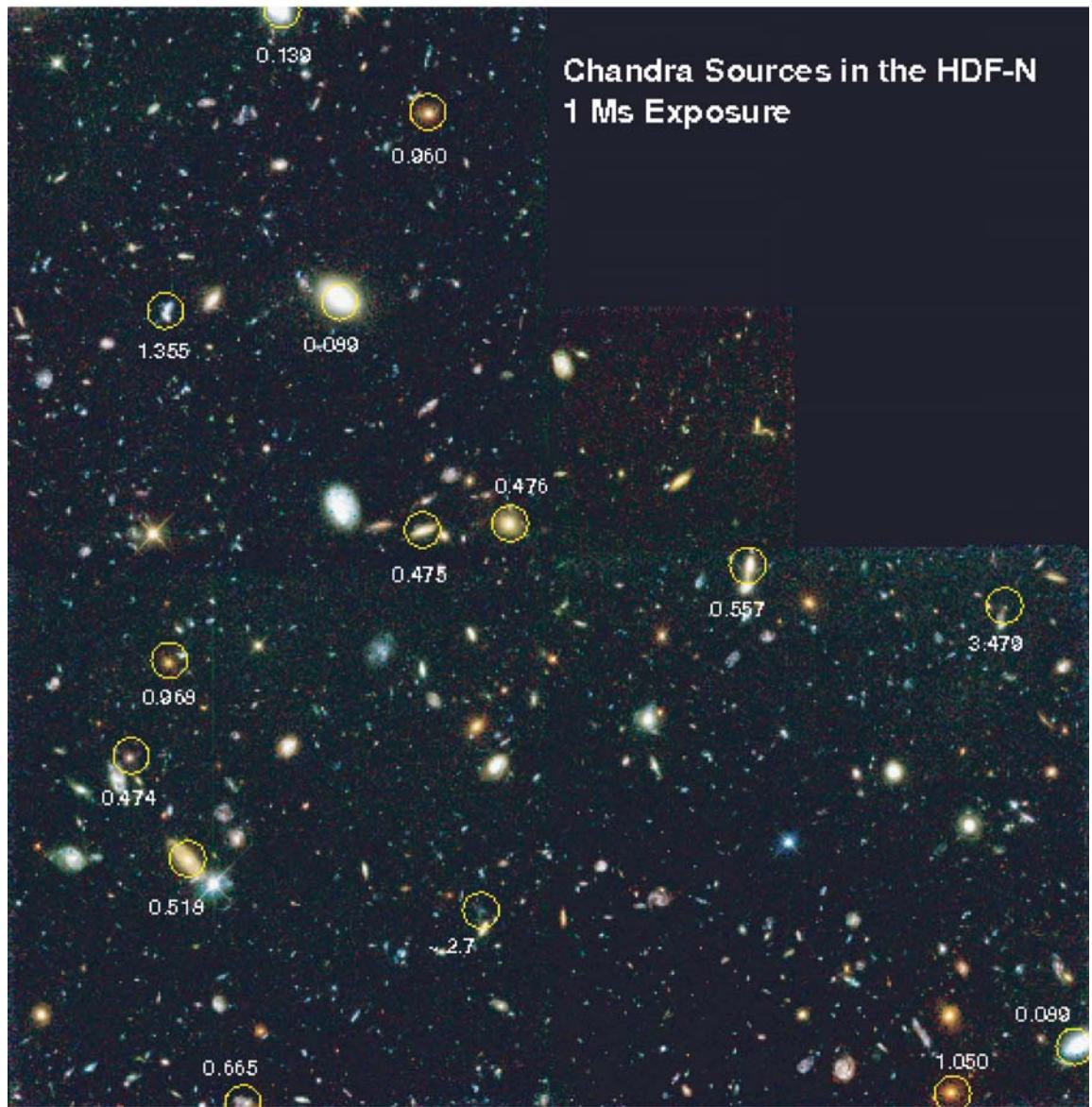


Plate 54 The 2-Ms Chandra Deep Field North. The field is $\approx 20'$ square, is centred on the HDF and contains ≈ 500 sources. Energy is colour coded: red, 0.5–2 keV; green, 2–4 keV; and blue, 4–8 keV. The diffuse red sources are clusters or groups of galaxies (from Alexander *et al.*, 2003).



Chandra Sources in the HDF-N
1 Ms Exposure

Plate 55 Chandra deep survey sources superposed on the HDF which is $\approx 2.5' \times 2.5'$. X-ray locations are much better than the circles drawn. Numbers are measured red shifts (from Brandt et al., 2001).

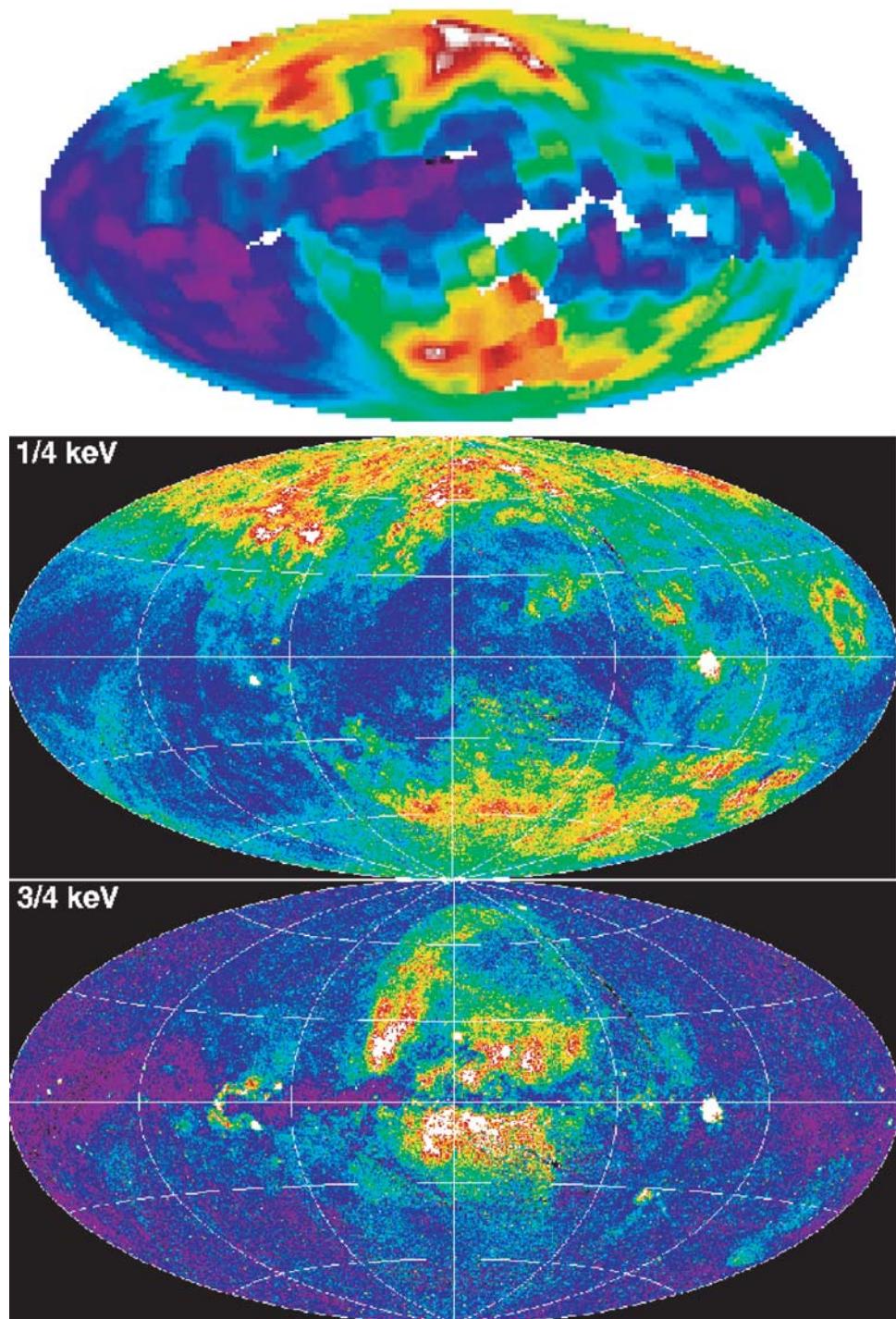


Plate 56 The soft X-ray background mapped by rocket and by ROSAT. Intensity ranges from blue (dim) to white (brightest). (top) Rocket results for the C band. The result is identical to the ROSAT data (middle) except for the improved resolution. (bottom) ROSAT M band map. Features are identified in Figure 16.7 (from McCammon et al., 1983; Snowden et al., 1995; MPE/ROSAT 1997).

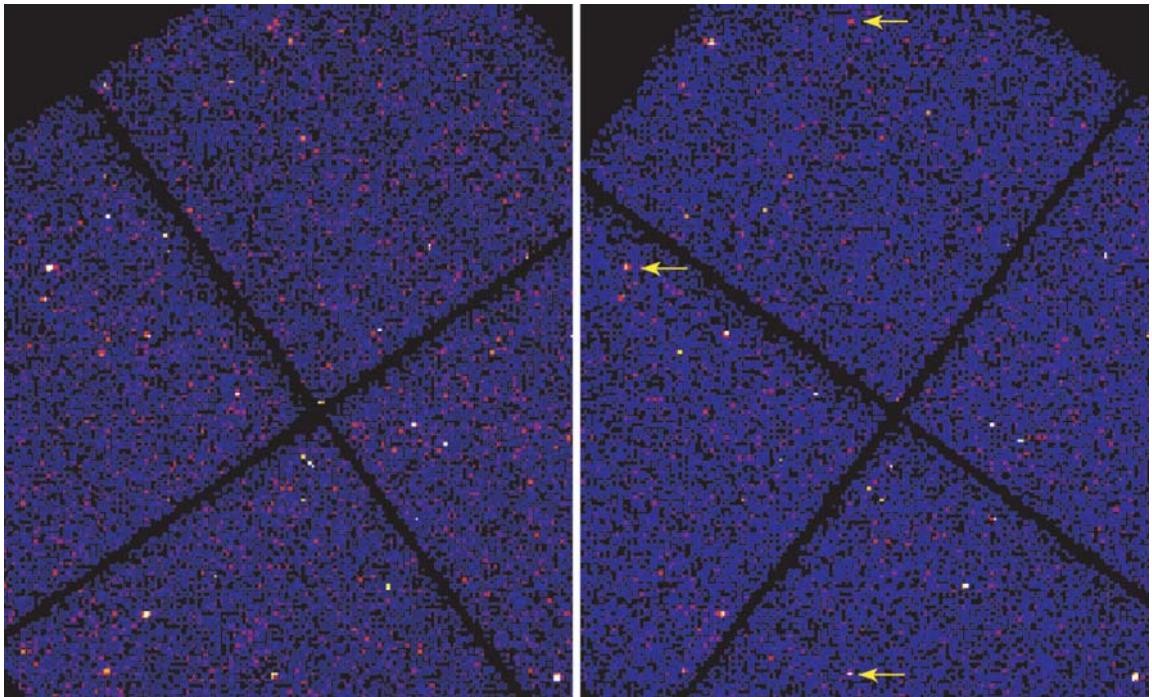


Plate 57 An attempt to identify the fading afterglow of XRF 040701. Two 3-hour Chandra ACIS observations taken 8 and 16 days after the outburst, were each centred on the HETE 8'-radius location. Fields shown are $12' \times 14'$ in extent. Three sources are indicated in the second observation which have decayed in intensity by a factor of ≈ 2 . The favoured candidate is the leftmost source (which is associated with a $z = 0.21$ galaxy). Without further evidence, this identification remains uncertain. This illustrates the difficulty of making an identification without a precise location and with a late start (adapted from Fox *et al.*, 2004).

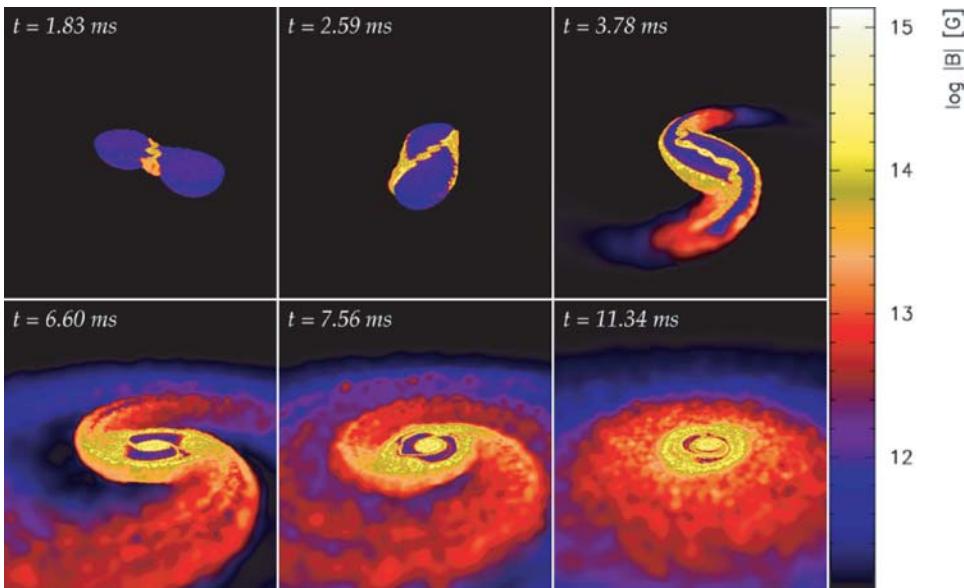


Plate 58 Final moments (tens of milliseconds) in the merger of two magnetised neutron stars. This computer simulation demonstrates the rapid formation of an accretion disc and strong magnetic fields. Colour indicates magnetic field strength. The initial neutron star field of 10^{12} G (blue) is amplified to strengths up to 10^{15} G (yellow), and this field propels the GRB material (from Price & Rosswog, 2006).

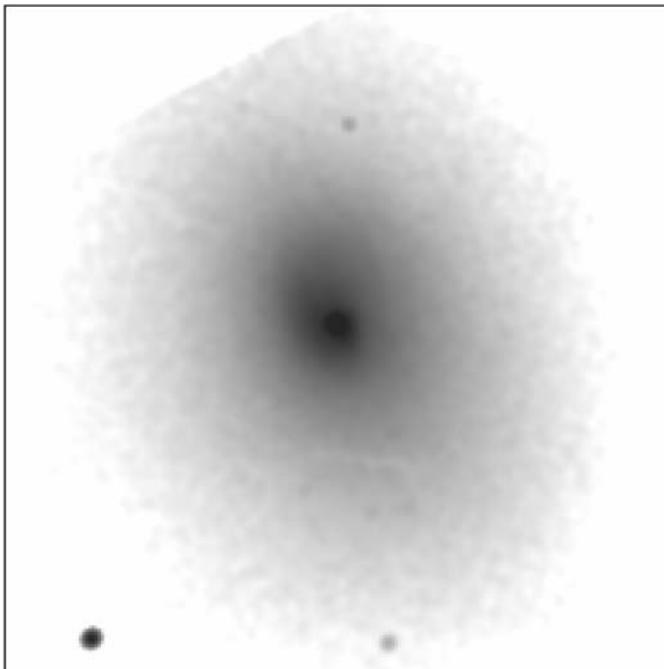


Fig 15.4 The cluster A2029 imaged in a 5-hour Chandra ACIS observation. The bordering box is 7' square – 0.68 Mpc at the cluster. Data have been smoothed and the grey scale is logarithmic. This is a regular, relaxed cluster with a cusp of soft emission at the location of a central cD galaxy. Note the gradual fading to background in all directions. The edge of the detector is seen in the NE. Emission from the ICM is slightly elliptical, which is not unusual (courtesy M. Markevitch).

X-ray-emitting ICM. Dark matter was more centrally concentrated than the ICM gas ($\beta = 0.75$). Some emission was detected from the vicinity of the large galaxies.

The increased sensitivity and better resolution of XMM resulted in the X-ray map in Colour Plate 48. The bright central region was then divided into ≈ 100 small areas and an accurate temperature was obtained for each (Colour Plate 47). Temperature in the central area around the two bright galaxies is almost constant and falls gradually with increasing distance from the centre. There is a cool area associated with NGC 4921 and hot spots associated with the in-falling NGC 4839 group, thought to be due to compression of the ICM. The Coma Cluster seems indeed relaxed in the central region, but the merging process is still ongoing in the outer parts. It is not quite the archetypical relaxed cluster.

The two central galaxies themselves now show as regions of relatively cool gas. A Chandra image shows X-ray emitting regions of 10^7 K gas at the positions of the two large galaxies (Vikhlinin *et al.*, 2001). These regions are considerably smaller than the optical sizes of NGC 4874 and NCG 4889 but larger than point sources. Cool gas associated

with these galaxies has apparently been compressed by the higher pressure of the surrounding 10^8 K ICM gas.

15.8 A2029: Rich, regular and relaxed, with cool centre and cD galaxy

A2029 is a rich cluster in Serpens containing thousands of galaxies and a large elliptical galaxy at the centre. The brightest galaxies in rich clusters are usually classified cD. The D galaxies are dustless systems dominated by amorphous light. The prefix c means supergiant. The cD galaxies are observed to have faint, extensive optical envelopes about a nuclear region, which sometimes has multiple components. Some, such as M87, are truly enormous. The X-ray surface brightness of A2029 is symmetrical and has no fine-scale structure other than a cusp at the centre at the location of the centre of the cD galaxy. Figure 15.4 shows a 5-hour Chandra image of the bright central region of the cluster. Faint emission can be traced to a radial distance of 10', which is 4 times farther than the

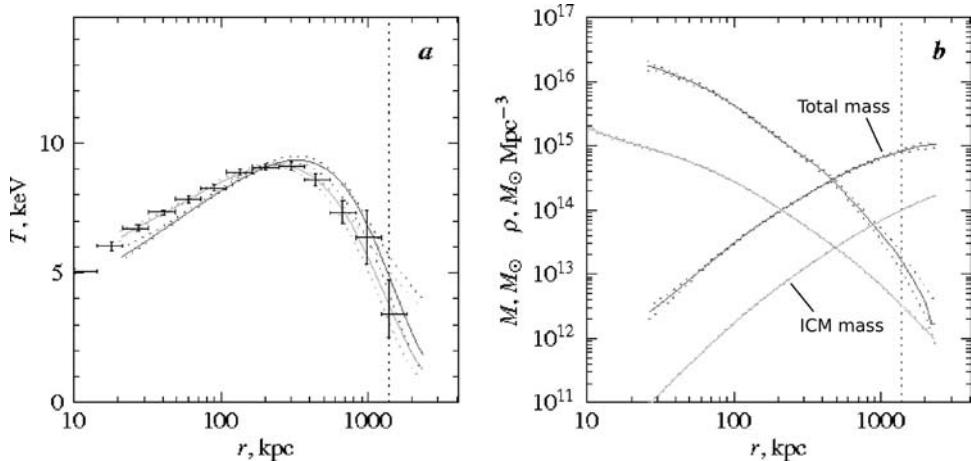


Fig 15.5 Radial profiles of (a) temperature and (b) mass distribution in A2029 from Chandra observations. Solid curves in Fig. 15.5a show three-dimensional model and corresponding projected temperature profile. Figure 15.5b shows mass density and total mass inside radius r . Heavy curves are total mass and mass density. Light curves are ICM (gas) mass and ICM density. Vertical line indicates r_{500} . Units of mas density are not intuitive: $10^{16} M_\odot \text{Mpc}^{-3}$ is $\sim 0.4 \text{ atoms cm}^{-3}$. The ICM gas density at the centre is $\sim 0.1 \text{ atoms cm}^{-3}$ and at the outer boundary is $\sim 10^{-4} \text{ atoms cm}^{-3}$ (from Vikhlinin et al., 2006).

apparent edge of the cluster in this figure. This is one of the best examples of a relaxed cluster. There is no sharp structure, but at the centre, something is happening.

Figure 15.5 shows the observed radial variation of temperature, mass and density. The cluster has a cool centre. Outside of this, the gas is not isothermal, but temperature drops as radius increases. The mass is calculated assuming hydrostatic equilibrium (as in Box 15.3) and using a modified beta model. As is the case for all clusters, $\beta \leq 1$ indicates that the ICM density falls off less rapidly than the density of dark matter. This can be seen in the A2029 analysis shown in Fig. 15.5. At the largest radius considered, the ICM accounts for 10 per cent of the total mass. Mass can only be determined out to the radius where cluster emission is obscured by the background. This sometimes results in underestimates for masses of distant faint clusters. When cluster masses are compared, they are generally truncated at a particular radius.

When the gas at the cluster centre has low temperature and high density, the radiative energy loss from X-ray emission can be large enough to have a significant effect on the behaviour of the gas. Energy is radiated, the gas cools and pressure drops. The gas is further compressed by gravity

Box 15.3 | Hydrostatic Equilibrium

To calculate cluster mass, spherical geometry is assumed. The mass inside radius r is $M(r)$. Pressure p , gas density ρ and temperature T all vary with r .

For pressure equilibrium,

$$dp/dr = -G M(r)\rho/r^2,$$

and for a perfect gas,

$$p = kT\rho/\mu m_p.$$

The mean molecular weight in AMU, μ , = 0.6, and m_p , is proton mass.

These combine to give

$$M(r) = -\frac{kTr}{G\mu m_p} \left[\frac{d \ln \rho}{d \ln r} + \frac{d \ln T}{d \ln r} \right].$$

Note the dependence on the temperature gradient, which can be hard to measure accurately.

and somewhat by pressure of the hotter surrounding gas. The gas sinks further into the cluster potential well. The density increases, and since the rate of radiation goes as the square of the density, energy is radiated still more rapidly. This continues until the temperature is low enough for the

Table 15.3 X-ray flux and luminosity for clusters in this chapter.

Cluster	Flux 2–10 keV (10^{-11} erg cm $^{-2}$ s $^{-1}$)	Red shift (z)	Distance ^a (Mpc)	L_X 2–10 keV (10^{44} erg s $^{-1}$)
A426 (Perseus)	75	0.0183	79	5.6
A1656 (Coma)	32	0.0235	101	3.8
M87 (Virgo)	30	0.0037	16	0.05
A2029	7.5	0.0767	335	10
A2142	7.5	0.0899	390	14
A2199	7.1	0.0305	132	1.5
A85	6.4	0.0518	225	3.8
A2256	5.2	0.0581	252	4.2
A1367	3.4	0.0215	93	0.36
1E 0657-56	1.1	0.296	1360	20

^a $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

gas to condense into filaments and presumably, eventually stars. This inward motion of the cluster gas is the ‘cooling flow’.

The cooling time of the hot ICM, with density n positive ions cm $^{-3}$ and temperature kT keV, radiating bremsstrahlung, is

$$t_{\text{cool}} \approx 2.5 \times 10^7 n^{-1} (kT)^{1/2} \text{ years.}$$

Gas density at the centre of A2029 (radius ≤ 10 kpc) is ≥ 0.06 atoms cm $^{-3}$, and temperature is 3 keV. The cooling time is $\approx 10^9$ years, clearly less than the $\approx 10^{10}$ year age of the cluster. The cooling flow in A2029 was originally calculated to produce $\sim 100 M_\odot$ of cold gas yr $^{-1}$, or $\sim 10^{11}$ stars in 10 9 years, enough to make a large galaxy. However, there is no observational evidence for this amount of cold gas. Neither excessive H α filaments nor star-forming regions have been detected. This cool-core-but-no-observed-cooling-flow puzzle occurs for many relaxed clusters.

It is thought that there must be a heating mechanism operating which in general prevents cooling flows from reaching the end of the process. There are several possibilities: (1) supernovae in a central galaxy, (2) conduction of heat from the surrounding hot gas, (3) in-falling groups could compress and heat the gas and (4) AGN activity in one of the cluster galaxies. All these processes may contribute, and not necessarily

in the same proportions in all clusters. Indeed, some clusters show spectacular disruptions of the ICM due to AGN outbursts (e.g. the Perseus Cluster).

The relaxed cluster, then, is the end point of cluster evolution (at an age of $\sim 10^{10}$ years), and straightforward techniques have been developed to determine mass and structure. The X-ray-emitting ICM is composed of primordial H and He and gas stripped from galaxies which contains metals made in stars. Abundances at the centre of A2029 are approximately solar and drop rapidly to ≈ 0.1 solar at large radii.

15.9 | A2142: Cold fronts

Figures 15.6 and 15.7 show a Chandra ACIS observation of the cluster A2142. This image shows two relatively sharp X-ray brightness edges, one south and one northeast of the cluster centre. Density and temperature profiles were extracted using regions contoured to follow the sharp edges. As one proceeds inwards, at the edges, the density increases and the temperature drops, instead of increasing as at a shock. In this case, the brightness edges do not indicate shocks and are called *cold fronts*. They are interpreted as contact discontinuities, boundaries between gas clouds and a warmer less-dense medium. The cooler clouds are moving

Fig 15.6 A 7-hour Chandra ACIS image of the cluster A2142. The bordering box is 8' square – 0.90 Mpc at the cluster. Data have been smoothed, and the grey scale is logarithmic. Two cold fronts are visible; one is the abrupt northern boundary where, in this figure, gray turns to white and one is just south of the centre where black turns to gray. The edge of the detector can be seen in the SE (courtesy of M. Markevitch).

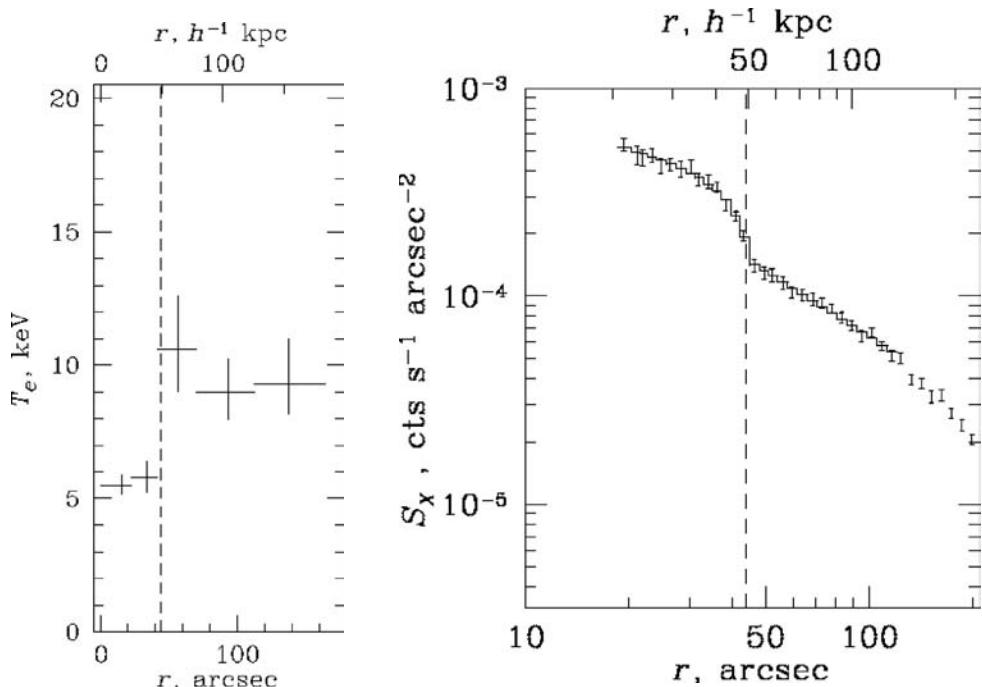
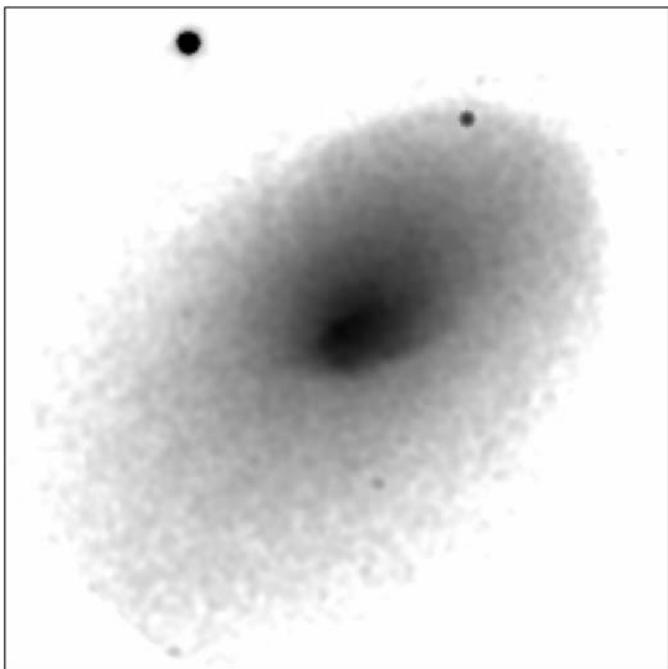


Fig 15.7 The southern cold front in A2142. Measured surface brightness and temperature profiles extending south from the cluster centre through the sharp decrease in brightness. Note the lower temperature in the high-density region (from Markevitch et al., 2000).

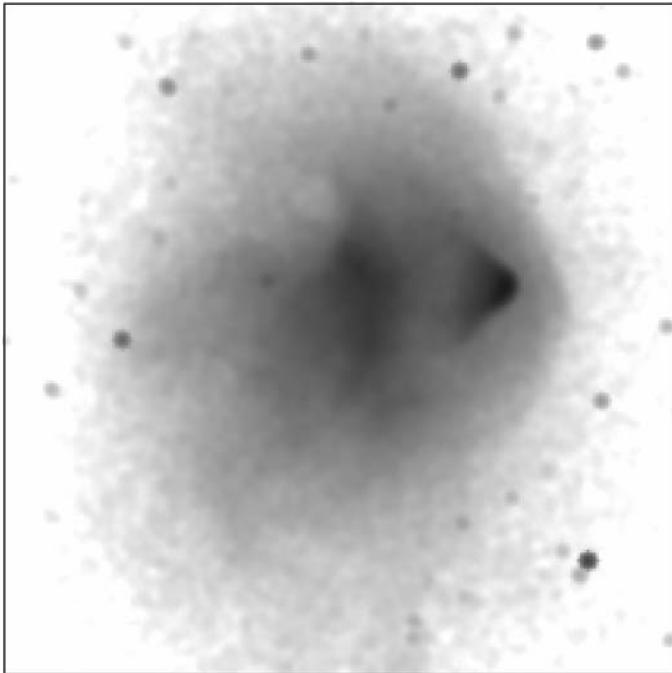


Fig 15.8 A 130-hour Chandra ACIS image of the Bullet Cluster. The bordering box is 7' square – 2.8 Mpc at the cluster. Data have been smoothed, and the grey scale is logarithmic. Maximum brightness comes from the westward-heading bullet, which is preceded by a shock. These data are also shown in Colour Plate 50 (courtesy of M. Markevitch).

subsonically through the warmer ICM. There is also pressure equilibrium across the density discontinuity, in contrast to the sudden increase in pressure that occurs at a shock front.

Many clusters show cold-front structures. They are probably the remnants of cool cores of the pre-merger clusters that have survived the impact and not yet mixed with the warmer gas. Some relaxed clusters also have rather subtle cold fronts within their cool cores. These are interpreted as a ‘sloshing’ of the dense, cool gas caused by a past cluster in-fall. Perhaps the most dramatic cold front is seen in the cluster 1E 0657-56, which is the topic of the next section.

15.10 The Bullet Cluster (1E 0657-56): A merger

Figure 15.8 and Colour Plate 50 show the Chandra X-ray image of the Bullet Cluster. In Colour Plate 50 this is superposed on an independent determination of the mass distribution. Two concentrations of dark matter separated by $\approx 4'$ have been mapped by gravitational lensing. This is an ongoing merger in the plane of the sky. The smaller

cluster has fallen in from the east and has passed through the major mass concentration. The ICM of the smaller cluster has been stripped by ram pressure, leaving a cool core travelling west and lagging behind the main mass of the subcluster. This cool core is cone shaped and shown in red in Colour Plate 50. The black-and-white image in Fig. 15.8 is printed to show the fainter emission. The cool bullet is preceded by a bow shock, which is a clear enhancement in the faint emission pervading the figure. The temperature profile in Fig. 15.9 shows that the cold front is preceded by a shock. The gas which forms the bullet was probably a cool core in the small cluster before the collision.

The fact that the dark matter distribution (blue in Colour Plate 50) is in two clumps after the collision implies that the dark matter is not fluidlike. Using these data, a limit can be placed on the cross section for dark matter self-interaction; the start of an interesting area of research.

15.11 M87 and the Virgo Cluster

The Virgo Cluster is the nearest large cluster and subtends an angle of more than 10° on the sky.

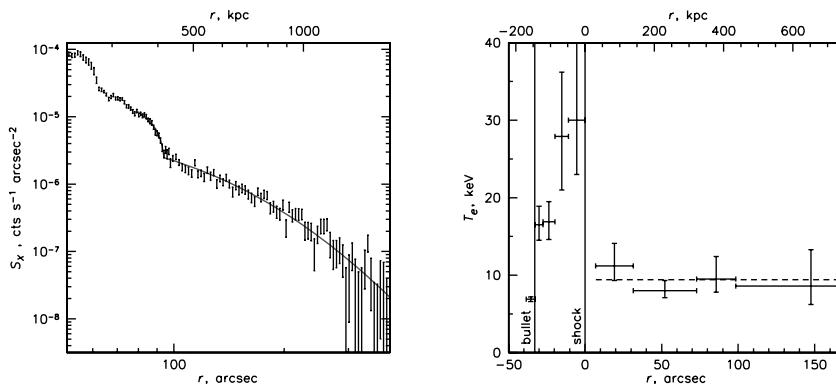


Fig 15.9 The Bullet Cluster: (left) an X-ray surface brightness profile starting within the bullet feature and proceeding westward; solid line shows a model with spherical density discontinuity; (right) a projected temperature profile across the bullet and shock; vertical lines mark boundaries of bullet and shock. Compare with Fig. 15.8 (from Markevitch *et al.*, 2006).

It is not relaxed. The most extensive map of X-rays from the ICM has been made by the ROSAT sky survey (Bohringer *et al.*, 1994) and is shown in Colour Plate 51. A halo centred on the dominant galaxy M87 accounts for 70 per cent of the diffuse X-rays, but the cluster is seen more as a collection of individual galaxies than as a definite structure. Other galaxies within the cluster can be seen as individual X-ray sources, but none approaches the strength of M87.

M87 contains a bright radio source with the famous optical jet extending about 30 arcsec from the nucleus. X-ray emission from this jet has been discussed in Chapter 14. The optical galaxy is very large. It has a measured size of 4 arcmin on the Palomar Sky Survey, and a faint optical halo extends at least 20 arcmin from the nucleus. Measured velocities of stars in the central region imply a mass of $\approx 2 \times 10^{11} M_\odot$ (equal to the mass of the stars in our entire Galaxy) within $1'$ (≈ 5 kpc) of the nucleus. The X-ray surface brightness implies a dark-matter mass of $\approx 3 \times 10^{13} M_\odot$ inside a radius of $50'$ (≈ 250 kpc). Thus M87 is 10 times as large as our Galaxy and over 100 times as massive. At this radius it is difficult to say how much dark matter is associated with M87 and how much with the cluster. Models used to fit brightness and temperature profiles of the outer region require an extra dark-matter component to account for M87, and this contributes ≈ 20 per cent of the mass within $50'$ (Nulsen & Bohringer, 1995). X-rays from the central few arcminutes show structure

associated with a radio halo. This is thought to be from gas cooler than the ICM and associated with the relativistic particles injected by the AGN of M87.

The great mass and the central position of M87 within the Virgo Cluster are not accidental. M87 was probably large when first formed and acquired more mass early in the life of the cluster. Even if it had not been formed at the centre of the cluster, collisions between galaxies will have settled this massive galaxy to the centre faster than the lighter ones. Such giant galaxies are only found in cluster centres, and as for M87, cluster and galaxy are intimately connected. The galaxies have grown to their present size by accretion of gas and by merging with smaller galaxies. They can only exist in clusters where conditions are right for such mass accretion. The process of the cD galaxies acquiring material from their smaller neighbours has produced the delightful expression ‘galactic cannibalism’: the giant galaxies tear their companions to bits with tidal forces and feed on the remains.

Other, smaller galaxies within the Virgo Cluster also show an interaction with the hot intra-cluster gas. A region centred on the galaxy M49, well south of M87, accounts for ≈ 10 per cent of the emission and ≈ 5 per cent of the mass and does not appear to be reacting strongly with M87.

The area around the galaxy M86, northwest of M87, accounts for ≈ 10 per cent of the diffuse X-rays. This is thought to be a group falling

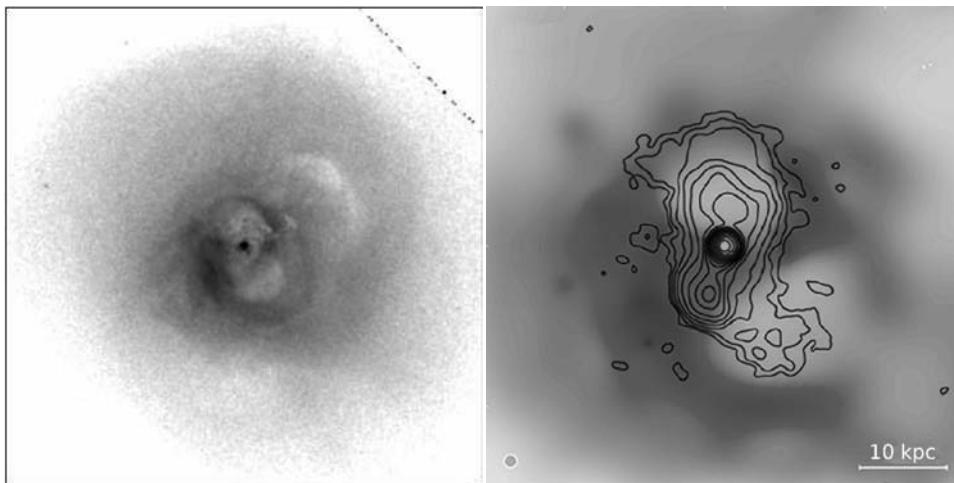


Fig 15.10 (left) A 27-hour Chandra ACIS observation of the centre of the Perseus cluster. The bordering box is 6' square – 140 kpc at the cluster. Data have been smoothed, and the grey scale is logarithmic. The detector edge is in the NW. Prominent features are the AGN at the centre, cavities north and south of this nucleus and a third cavity in the NW (courtesy of W. Forman), (right) Perseus Cluster X-ray image with contours of constant radio surface brightness superposed. The radio contours fill the two cavities north and south of the nucleus. The field is 2.5' square – 57 kpc at the cluster (NASA/CXC/IoA 2003; Fabian et al., 2001).

towards M87 with a relative velocity of $\approx 1500 \text{ km s}^{-1}$. An $\sim 60\text{-ks}$ XMM observation of M86 shows the interaction with M87 to be complex (Finoguenov *et al.*, 2004). The M86 group appears intact and almost relaxed. The temperature structure shows a shock where the group is colliding with the medium around M87, but outside the virial radius of M87. Gas around the M86 galaxy has been disturbed, probably by collisions with other galaxies rather than with the M87 ICM. Thus the nearest large cluster, which is not relaxed, shows some details of the merging process.

15.12 The Perseus Cluster: AGN, cavities and sound waves

The Perseus Cluster, A478, is the brightest X-ray cluster in the sky. The active galaxy NGC 1275 is at the centre, and the ICM shows evidence that the black hole at the centre has been periodically depositing energy into the surrounding gas. This deposition is large enough to offset radiative cooling.

There are two cavities in the ICM north and south of the central galaxy. Figure 15.10 overlays

X-ray and radio images and shows that the X-ray cavities are filled with radio-emitting relativistic particles. In the past, accretion onto the black hole has produced jets which have deposited $\approx 4 \times 10^{58}$ ergs into each cavity. In the future, these energy bubbles will rise buoyantly and travel outwards. A large cavity to the northwest, without radio emission, no longer contains relativistic plasma and is probably the remnant of a past event – a ‘ghost bubble’. The rising bubbles push against and compress the ICM gas. This forms a shock which can be seen in Colour Plate 52 as a bright arc with sharp boundary at the top of the inner northern bubble. The shock is recognised by an increase of pressure and density within the arc. The gas temperature, however, is observed to be the same as that of the surrounding gas. Apparently thermal conduction of the ICM is high. Farther out in Colour Plate 52, ripples in the ICM can be seen. These are emphasised in Fig. 15.11, which was produced by applying an unsharp mask to the X-ray image. The image smoothed with a $2''$ Gaussian filter was divided by the image smoothed with a $20''$ filter. In this way, large-scale structure was eliminated and the result shows the arcsecond ripples clearly. Pressure, density and temperature indicate that these nearly concentric

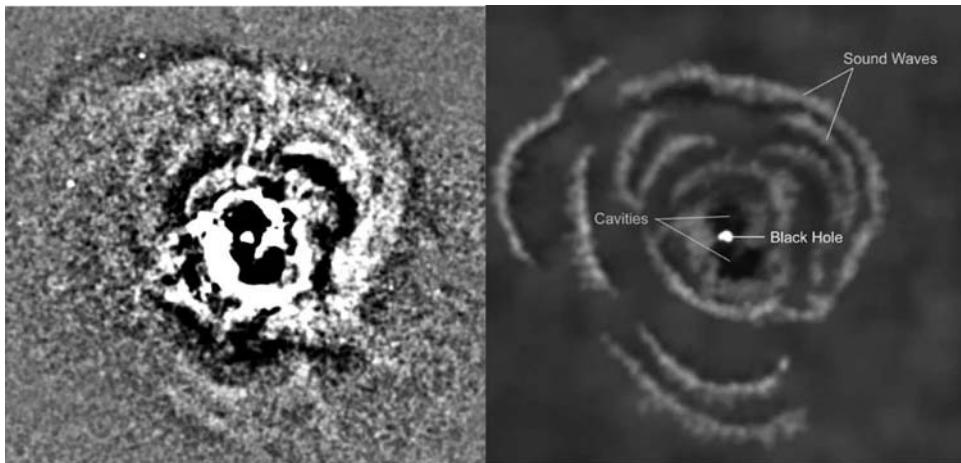


Fig 15.11 (left) Perseus Cluster image processed with an unsharp mask to show structure. The outer ripples are now clear as is structure around the nucleus. (right) Artist's drawing of ripples with features labelled (NASA/CXC/LoA 2003; Fabian et al., 2006).

ripples are sound waves. The regular spacing and a sound speed of $\approx 10^3 \text{ km s}^{-1}$ imply outbursts every $\sim 10^7$ years.

Here is evidence for a cyclic process. A cooling flow in the ICM produces cold gas and stars which sink to the central region of NGC 1275. The black hole captures some of this material. Increased accretion leads to jet formation, which pumps energy into the ICM. The transfer mechanism is through sound waves. This heating turns off the cooling flow and accretion. Power in the jet decreases and remains low until the ICM cools and the cycle repeats. Thus cooling flows can be turned off or moderated (Fabian et al., 2006; Tucker et al., 2007). Other clusters also show signs of this process. The largest deposition of energy observed so far is in the distant cluster MS 0735.6+7421. Cavities on two sides of the nucleus imply injection of $\approx 6 \times 10^{61}$ ergs (Gitti et al., 2007).

15.13 The relation of mass, temperature and X-ray luminosity

The luminosity and temperature of the ICM depend on the mass of the cluster. The deeper the gravitational potential, the more hot and dense the ICM, which means that L_X , M and T are

related. Bigger clusters are hotter and brighter than small ones. Since all the larger relaxed clusters appear to have the same structure, the following considerations should apply.

The temperature should scale with the gravitational potential,

$$T \propto M/R,$$

and the mass should scale with volume,

$$M \propto R^3 \text{ or } R \propto M^{1/3}.$$

Thus we expect

$$M \propto T^{3/2} \text{ and } T \propto M^{2/3}.$$

And for bremsstrahlung spectra,

$$L_X \propto \rho^2 R^3 T^{1/2} \propto M^{4/3}.$$

Figures 15.12 and 15.13 show observed correlations of L_X and T with M for nearby relaxed clusters. In these figures the mass shown is the mass included within R_{500} (inside the virial radius and with good X-ray signal to background ratio) but with the central 15 per cent excluded to avoid cooling effects. T is the average temperature in this region. L_X is measured in the ROSAT band-pass (0.5–2 keV). The measured slope of the logM-logT relation is 1.53 ± 0.08 , very close to the expected 1.50. The measured slope of the log L_X -logM relation is 1.61 ± 0.14 , steeper than the 1.33 expected (see Vikhlinin et al., 2009a, for details).

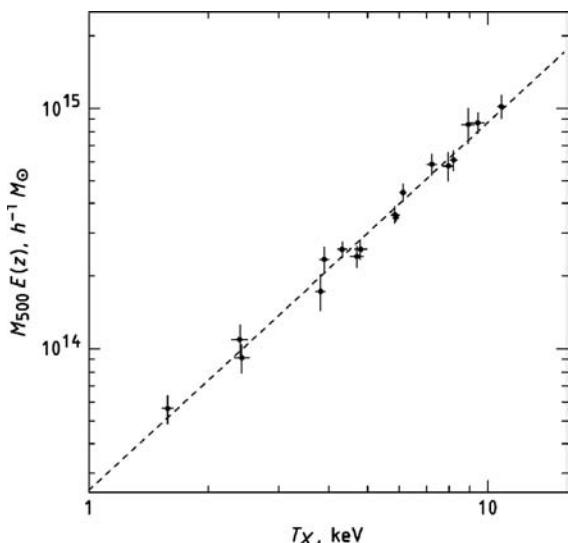


Fig 15.12 The relation between gravitational mass and gas temperature for 17 low- z , relaxed Chandra clusters. The dashed line is a best-fit power law with index 1.53. The factor $E(z) = H(z)/H_0$ is not large for these data (from Vikhlinin *et al.*, 2009a).

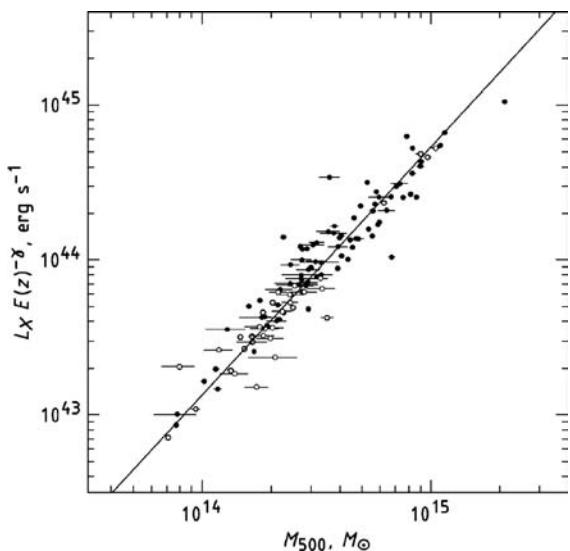


Fig 15.13 The relation between cluster mass and X-ray luminosity in the 0.5–2 keV band. Solid points are low- z clusters, and points with error bars are high- z clusters with luminosities corrected for evolution. Open points have large corrections. The solid line is the best-fit power law to the low- z data and has index = 1.61 (from Vikhlinin *et al.*, 2009a).

Using the measured values, the relations are $L_x \propto M^{1.6}$ and $L_x \propto T^{2.4}$ for relaxed clusters. For even smaller masses, the ICM of groups is known to be cooler than that of clusters. Mulchaey (2000) has shown that for groups, $L_x \propto T^5$. Since most groups are irregular with non-uniform temperature distribution, the interpretation of this result is not straightforward.

15.14 Distant clusters

The only way of ‘seeing’ cluster evolution is to study distant clusters. As red shift (z) increases, the light comes from further in the past. Since it is difficult to detect faint galaxies optically and there is confusion from overlapping clusters, X-ray observations have become the primary technique for identification of distant clusters. The hot ICM is a sure indication of the dark matter concentration in a cluster. The first X-ray detection of distant clusters was by the Einstein Observatory. Early in the mission, 11 clusters with z from 0.1 to 0.8 were detected and found to have size and luminosity ‘similar’ to nearby clusters (Henry *et al.*, 1979).

At early times, however, clusters are expected to be less X-ray luminous since, as a cluster evolves, both the gas content and the central potential increase. This has now been confirmed by studies which have examined enough clusters to be significant. For example, the 160-square-degree ROSAT survey (Mullis *et al.*, 2003) used a sample of 201 clusters. Distant clusters ($0.3 < z < 0.8$) were compared to the X-ray luminosity function of local clusters. There was agreement between the two groups at low luminosity, but for $L_x > 3 \times 10^{44}$ ergs s^{-1} , only three clusters were found in the high- z sample, where 15 were expected. Because mass and luminosity are well correlated (Fig. 15.13), this implies a deficit of high-mass, or evolved, clusters at early times.

Furthermore, cosmology is important. The gravitational collapse of a cluster starts early in the life of the Universe and captures a sample of primordial material. Knowing the cluster mass function (the integral number of clusters as a function of mass), the average density of matter can be calculated. This was done by Reiprich

Box 15.4 Cosmology: The critical density

The dynamics of the Universe are described by the Friedman equations:

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3}G\rho - \frac{kc^2}{a^2} + \frac{\Lambda}{3}$$

$$\frac{\ddot{a}}{a} = \frac{\Lambda}{3} - \frac{4\pi}{3}G\left(\rho + \frac{3p}{c^2}\right).$$

H is the Hubble constant (which has value H_0 at the present time), a is a scale factor (the size of a piece of space), G is the gravitational constant, k is the curvature parameter, ρ is density, p is pressure and Λ is the cosmological constant (representing the energy density of empty space).

Assuming $k = 0$ (the Universe is flat) and $\Lambda = 0$, then the density required to just stop the present expansion of space is

$$\rho_c = \frac{3H_0^2}{8\pi G}.$$

This is the *critical* or *closure* density. At present, $\rho_c = 0.947 \times 10^{-26} \text{ kg m}^{-3}$, the equivalent of $\approx 6 \text{ H atoms m}^{-3}$.

The normalised density parameter, Ω , is density expressed in units of the critical density. In the present Λ -cold-dark-matter (Λ CDM) model of the Universe, the density of matter, $\Omega_{m,0} = 0.27 \pm 0.04$ (of which 0.23 is dark matter and 0.044 ± 0.004 is stars and hot gas), the density of radiation and neutrinos, $\Omega_{\text{rel},0} = 0.00008$, and the density of dark energy, $\Omega_{\Lambda,0} = 0.73 \pm 0.04$. The Hubble constant is $H_0 = 71 \pm 4 \text{ km s}^{-1} \text{ Mpc}^{-1}$, and the age of the Universe is 13.77×10^9 years.

Each entity also has an equation of state, expressed as w , the ratio of pressure to density. Accepted values are $w_m = 0$ (for cold dark matter), $w_{\text{rel}} = 1/3$ and $w_{\Lambda} = -1.00$.

and Böhringer (2002) using a ROSAT determination of the dependence of L_x on M . Their early result, $\Omega_m = 0.12 \pm 0.5$, was about half the generally accepted value (see Box 15.4).

A more recent work by Vikhlinin *et al.* (2009a, 2009b) has achieved a good demonstration of the effect of cosmology on derived early cluster properties. Their sample comprised 36 high-red shift

clusters from the ROSAT 400-square-degree survey and 19 bright clusters from the ROSAT all-sky survey. All these were again observed with Chandra and are indeed relaxed, regular clusters. The mass used, M_{500} , was taken from a model fit to both ROSAT and Chandra data in the radial region $(0.15–1.00)R_{500}$. Because of the possibility of cooling or emission from central galaxies, central regions were excluded. The cluster mass functions are shown in Fig. 15.14, which clearly shows cluster evolution. There were fewer large clusters at early times. In addition, dark energy is required to fit the subset of high- z clusters shown. (Please realise that this analysis was complex, and we only skim the results here.) Note that the scales in the figures contain the factor h^{-1} , which shows the dependence of distance calculated using a particular value of the Hubble constant. The particular cosmology model used affects the high- z cluster data points. If there were no dark energy, the high- z clusters would be less massive and occupy a smaller volume of space. This data set requires the existence of dark energy, $\Omega_{\Lambda} > 0$, and measures the dark energy equation-of-state parameter $w = -1.14 \pm 0.21$. The expected Λ CDM value is -1.00 .

15.15 The Sunyaev-Zeldovitch (SZ) effect

In 1978, at the first Einstein review of Guest-Observer proposals, the two most interesting proposals – those which provoked the most lively panel discussion – were almost identical. Both proposed ~ 10 -hour observations of the cluster A2218. These X-ray data, a microwave spectrum, and the SZ effect would be used to measure the Hubble constant to an accuracy of 10 per cent. Both proposals were deemed excellent, but one proposer showed a disdain for including quantitative information (e.g. angular size of the target) which might be used to evaluate feasibility. The more practical-sounding proposal was selected and the observation accomplished. As you might suspect, the project turned out to be considerably more complex than imagined in 1978. Now, 30 years

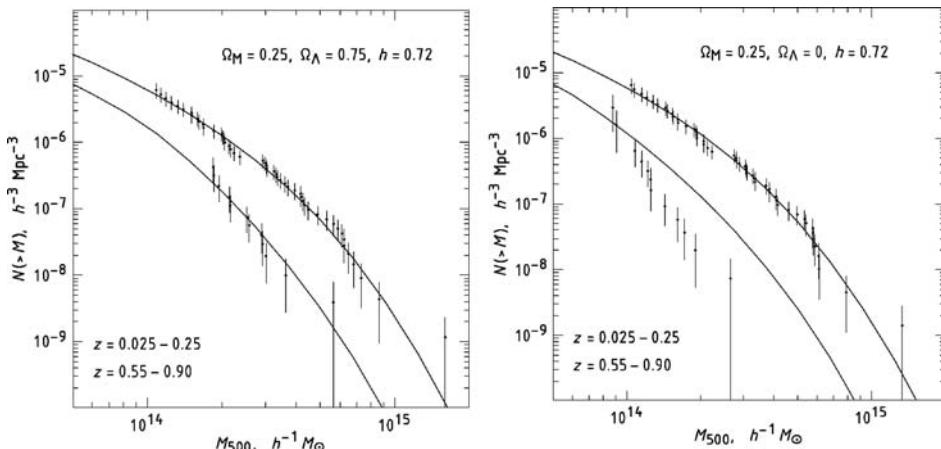


Fig 15.14 Cluster mass functions for low- and high- z samples. This shows the integral number of clusters with mass greater than a given mass. Vertical scale is number of clusters per megaparsec cubed. Solid curves show mass functions calculated for early and recent times. The two panels show two different cosmologies: (left) the presently accepted model with $\Omega_{\Lambda,0} = 0.75$ and (right) a model with no dark energy, $\Omega_{\Lambda,0} = 0.00$, which does not fit the observations (Vikhlinin et al., 2009b).

and hundreds of observations later, we are starting to see interesting results.

The SZ effect occurs when the 2.7-K cosmic microwave background is viewed through the centre of a regular cluster. The microwaves are Compton scattered by hot electrons in the ICM, and the peak of the microwave spectrum is shifted to a higher frequency. An accurate measure of the microwave intensity will give the thickness of ICM in the path. One must know accurately the ICM temperature, density (which depends on distance) and shape of the cluster core. The distance to the cluster can then be derived. This is not as easy as it sounds because most clusters are not regular, and the radio measurements are very difficult. Using Chandra observations of 38 clusters, Bonamente *et al.* (2006) derived $H_0 = 77 \pm 12 \text{ km s}^{-1} \text{ Mpc}^{-1}$. This demonstrates that the techniques employed are valid, but refinement is needed to match the precision of the generally accepted WMAP value of $H_0 = 71 \pm 4 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

References

Observation and theory

Arnaud, M., *et al.* (2001) XMM-Newton observation of the Coma Cluster, the temperature structure in the central region, *A&A* **365**, L67.

- Böhringer, H., *et al.* (1994) The structure of the Virgo Cluster of galaxies from ROSAT X-ray images, *Nature* **368**, 828.
- Bonamente, M., *et al.* (2006) Determination of the cosmic distance scale from Sunyaev-Zel'dovich effect and Chandra X-ray measurements of high-redshift galaxy clusters, *ApJ* **647**, 25.
- Briel, U. G., J. P. Henry & H. Böhringer (1992) Observation of the Coma Cluster of galaxies with ROSAT during the All-Sky Survey, *A&A* **259**, L31.
- Briel, U. G., *et al.* (2001) A mosaic of the Coma Cluster of galaxies with XMM-Newton, *A&A* **365**, L60.
- CXC (2007) <http://chandra.harvard.edu>.
- Fabian, A. C., *et al.* (2001) Chandra imaging of the complex X-ray core of the Perseus cluster, *MNRAS* **321**, L33.
- Fabian, A. C., *et al.* (2006) A very deep Chandra observation of the Perseus cluster: Shocks, ripples and conduction, *MNRAS* **366**, 417.
- Finoguenov, A. (2004) XMM-Newton witness of M86 X-ray metamorphosis, *A&A* **415**, 415.
- Forman, W., *et al.* (1972) Observations of the extended X-ray sources in the Perseus and Coma clusters from Uhuru, *ApJ* **178**, 309.
- Friedman, H., & E. Byram (1967) X-rays from sources 3C 273 and M 87, *Science* **158**, 257.
- Fritz, G., *et al.* (1971) Discovery of an X-ray source in Perseus, *ApJ* **164**, L81.
- Gitti, M., *et al.* (2007) XMM-Newton view of MS 0735+7421: The most energetic AGN outburst in a galaxy cluster, in *Heating vs. Cooling in Galaxies and Clusters of Galaxies*, 90, Garching, Germany: Springer.

- Gursky, H., et al. (1971) A strong X-ray source in the Coma Cluster observed by Uhuru, *ApJ* **167**, L81.
- Henry, J. P., et al. (1979) Detection of X-ray emission from distant clusters of galaxies, *ApJ* **234**, L15.
- Kellogg, E., et al. (1973) Clusters of galaxies with a wide range of X-ray luminosities, *ApJ* **185**, L13.
- King, I. (1966) The structure of star clusters. III. Some simple dynamical models *AJ* **71**, 64.
- Markevitch, M. (2006) Chandra observation of the most interesting cluster in the universe, in *The X-ray Universe 2005*, 723, Madrid, Spain: El Escorol.
- Markevitch, M., et al. (2000) Chandra observation of Abell 2142: Survival of dense subcluster cores in a merger, *ApJ* **541**, 542.
- Meekins, J., et al. (1971) X-rays from the Coma Cluster of galaxies, *Nature* **233**, 107.
- Mitchell, R. J., et al. (1976) Ariel 5 observations of the X-ray spectrum of the Perseus Cluster, *MNRAS* **175**, 29.
- Mullis, C. R., et al. (2003) The 160 deg² ROSAT survey revised catalog and cluster evolution, arXiv:astro-ph/0303438.
- NASA/CXC (2006) <http://chandra.harvard.edu/photo/2006/le0657-56>.
- NASA/CXC (2005) <http://chandra.harvard.edu/photo/2005/perseus>.
- NASA/CXC/IoA (2003) <http://chandra.harvard.edu/photo/2003/perseus>.
- Nulsen, P. E. J. & H. Bohringer (1995) A ROSAT determination of the mass of the central Virgo Cluster, *MNRAS* **274**, 1093.
- Reiprich, T. & H. Böhringer, (2002) The mass function of an X-ray flux-limited sample of galaxy clusters, *ApJ* **567**, 716.
- Vikhlinin, A., et al. (2001) Zooming in on the Coma Cluster with Chandra: Compressed warm gas in the brightest cluster galaxies, *ApJ* **555**, L87.
- Vikhlinin, A., et al. (2006) Chandra sample of nearby relaxed galaxy clusters: Mass, gas fraction, and mass-temperature relation, *ApJ* **640**, 691.
- Vikhlinin, A., et al. (2009a) Chandra Cluster Cosmology Project. II. Samples and X-ray data reduction, *ApJ* **692**, 1033.
- Vikhlinin, A., et al. (2009b) Chandra Cluster Cosmology Project. III. Cosmological parameter constraints, *ApJ* **692**, 1060.
- Vrtilek, J. (2007) private communication.

Catalogues, reviews, popular

- Abell, G. (1958) The distribution of rich clusters of galaxies, *ApJ Supp* **3**, 211.
- Fabian, A. (1994) Cooling flows in clusters of galaxies, *ARA&A* **32**, 277.
- Forman, W. & C. Jones. (1982) X-ray imaging observations of clusters of galaxies, *ARA&A* **20**, 547.
- Mulchaey, J. (2000) X-ray properties of groups of galaxies, *ARA&A* **38**, 289.
- Sarazin, C. L. (1988) X-ray Emissions from Clusters of Galaxies, Cambridge: Cambridge University Press.
- Tucker, W., H. Tananbaum & A. Fabian (2007) Black Hole Blowback, *Sci Am.* **296**, 42.
- de Vaucouleurs, G. (1975) Nearby groups of galaxies, *Stars Stell Syst* **9**, 577.
- Zwicky, F., F. Herzog & P. Wild (1966) *Catalogue of Galaxies and Clusters of Galaxies*, Pasadena: California Institute of Technology.

Chapter 16

The diffuse X-ray background

16.1 Early observations, 1965

The X-ray background was not anticipated. It was discovered in 1962 during the rocket flight which first detected Sco X-1, the first successful attempt to detect X-rays from sources other than the Sun or Earth. An uncollimated detector viewing about 10 000 square degrees of the sky was used. Giacconi *et al.* (1962) concluded that the background was of ‘diffuse character’ and due to X-rays of about the same energy as those from Sco X-1. The observed diffuse signal in this detector could have been generated by a few moderately strong point sources spread over the sky. The next observations, however, with detectors collimated to observe only 100 square degrees, showed the background to be indeed diffuse and of uniform brightness to at least 10 per cent.

There was no doubt that this background existed. The signals observed were strong and unmistakable. When detectors in rocket payloads were uncovered, pointed at any part of the night sky, the count rate always increased. All early observations, without exception, showed a few bright sources embedded in a uniform X-ray glow. The night sky at X-ray wavelengths was uniformly bright! Sources appeared superposed on this background, rather like stars viewed with the naked eye on a night with a full moon; when the faint stars disappear into the background of moonlight scattered from the atmosphere. Because no structure was observed and the emission was apparently uniform, this phenomenon has been called the ‘diffuse X-ray background’. Attempts

to understand it first relied on the generation of precise maps using detectors having 10–100 square degree fields of view and on measurements of the background spectrum over a large energy range.

Early speculation centred on diffuse origins and was strongly influenced by observations at other wavelengths. Figure 16.1 shows the isotropic sky flux as a function of frequency from the radio band to γ -ray energies. The dominant feature, containing 95 per cent of the energy, is in the microwave-far-infrared band and has the spectrum of a blackbody at a temperature of 2.7 K. The optical and UV points shown are upper limits. Our Galaxy is opaque to UV radiation, and local conditions make it difficult to detect faint diffuse optical radiation. It is, however, transparent to X-rays, except for observations within a few degrees of the galactic plane.

Since above 1–2 keV, the Universe (except for clumps of cold material in galaxies) is now transparent, the discovery of an isotropic X-ray flux was exciting. Perhaps here was a high-energy phenomenon somehow analogous to the microwave background. Whether truly diffuse or the sum of faint individual sources, the high degree of isotropy implied that the source of radiation was at a great distance. The brightness indicated that this was not a rare phenomenon. The fact that the photons are X-rays implied a high-energy process, certainly much higher than the energy of hydrogen recombination which marks the origin of the microwave background. Two possibilities considered for a diffuse cosmic component were a tenuous hot gas filling the space between clusters

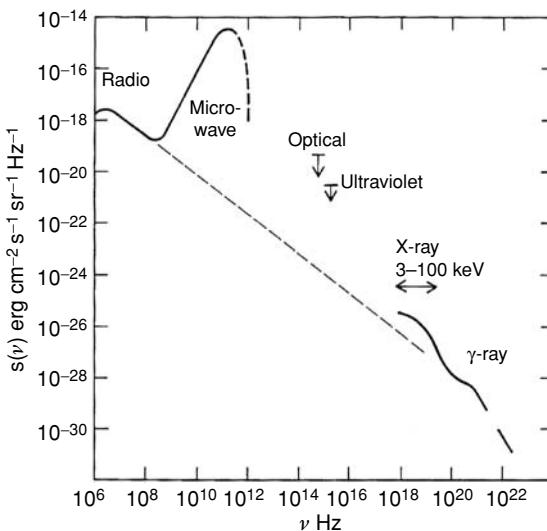


Fig 16.1 The isotropic background radiation as a function of frequency. The dashed line connects the radio and γ -ray bands and is a power law with index -0.7 . Ninety-five per cent of the energy is contained in the microwave band (courtesy R. Shafer).

of galaxies and inverse Compton photons from cosmic rays interacting with the microwave background.

We know now that this is not the case. The development of high-spatial-resolution telescopes and detectors in ~ 1980 gave the capability to find faint sources. The diffuse extragalactic background has now been almost completely resolved into a finite number of faint sources! However, it has not been conclusively demonstrated that 100 per cent of the background is due to discrete sources. A small fraction of truly diffuse emission of cosmic origin is still a possibility. Since this would imply a vast amount of previously unseen matter in the Universe, this is still seriously considered. Interest now, however, is centred on the ensemble of faint sources which comprise most, probably all, the background. These produce a strong signal and contain information about X-ray emission from the Universe at early times. In this chapter, we will review a bit of history, show some observations and summarise current knowledge about potential contributor sources. We then cover galactic emission which does indeed have a diffuse component.

Box 16.1 | The microwave background

The microwave background was discovered and explained in 1965 (Penzias & Wilson, 1965; Dicke *et al.*, 1965). It is universally accepted as ‘relic radiation from the big bang’ and is the most convincing bit of evidence that, in the past, matter in the Universe was much denser and hotter than it is now – a gross understatement. The ‘big bang’ theory postulates that the entire observable Universe originated in a configuration of extremely high density and temperature. All matter and energy first appeared in a soup of elementary particles. Since then, the density and temperature have been steadily decreasing as the Universe expands.

At an age of about 1 million years, particles that make matter as we know it had been formed. The Universe at this time was a sea of ionised material, mostly electrons, protons and He nuclei – a plasma, opaque to electromagnetic radiation. Photons could only travel short distances before colliding with free electrons. As the temperature decreased, photon energy decreased, and soon photons were not energetic enough to remove electrons from bound states with the protons. At this time, hydrogen was formed, and almost all electrons were captured and locked in hydrogen atoms. Since hydrogen is transparent, the photons were suddenly able to stream through space for great distances.

We are now able to look through space back to the time when the Universe was opaque. The photons we see are the UV-optical radiation freed by recombination but red-shifted by a factor of 1000 into the microwave range.

16.2 | HEAO-1, 3–30 keV, 1977

16.2.1 | A background map

The satellite HEAO-1, the first High Energy Astronomy Observatory, was launched in August 1977 and spent 3 years mapping the X-ray sky. It carried three sets of detectors, all with fields of view restricted by mechanical collimators: one set to detect discrete sources, one set to map the background and one set to explore the sky at high

energies. The sky was scanned by a slow rotation of the observatory.

One array of detectors was designed to determine the characteristics of the X-ray background. Six detectors with 400-cm^2 -area windows provided coverage in three energy bands from 0.1 to 60 keV. Special care was taken to reject internal and charged-particle induced backgrounds with internal anticoincidence requirements (as discussed in Chapter 3). Each detector had a collimator which defined two fields of view, one having twice the solid angle of the other. The diffuse signal was twice as great in the detector chambers viewing through the large-aperture part of the collimator. The internal background, however, was the same as in chambers viewing through the narrow aperture. Thus Elihu Boldt and colleagues (Rothschild *et al.*, 1979) attacked the instrumental problems which usually limit the accuracy with which diffuse signals can be measured. The background map produced from observations with this instrument is still the best available in this energy range.

In 3 years of operation, HEAO-1 performed 2.7 complete scans of the sky. Coverage was complete, but more time was spent looking at the ecliptic poles than at the ecliptic plane. Rotation of the spacecraft fixed the scan path to be a great circle perpendicular to the Earth-Sun line. This path crossed the poles once every rotation but precessed over the rest of the sky once every 6 months.

Data were used to make an all-sky X-ray map (Colour Plate 53) with $3^\circ \times 3^\circ$ and $3^\circ \times 6^\circ$ collimators. The map is dominated by bright sources clustered in the central region of the galactic plane. The first step in the study of the background was to remove the effect of these sources. This was not difficult since the response of the detectors to a pointlike source was well calibrated and source strengths and locations were accurately determined. After making a catalogue of detected sources (shown in Colour Plate 2), the signal due to these was subtracted. The ‘source-free’ map shown in Colour Plate 53 was the result. To improve counting statistics, the data here have been binned into $9^\circ \times 9^\circ$ degree sections of the sky. Almost all of the events comprising this map are due to the extragalactic X-ray background.

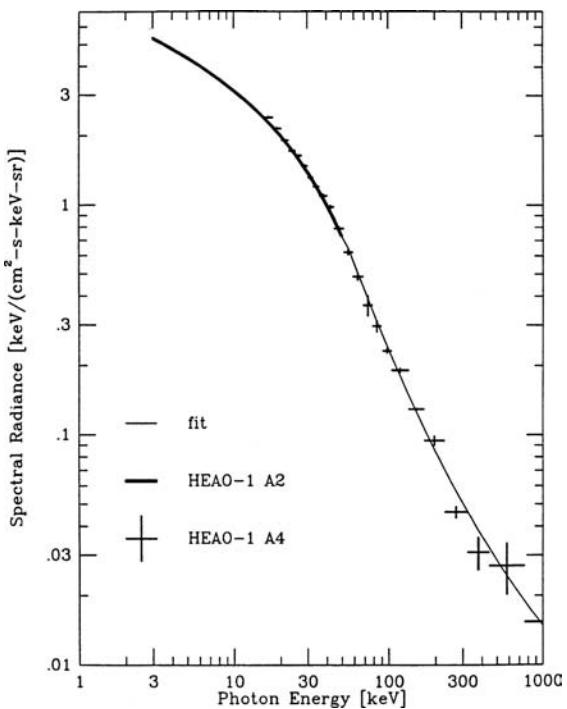


Fig 16.2 The spectrum of the X-ray background. The solid line from 3 to 50 keV is from HEAO-1 proportional counters. The accuracy of this spectrum is remarkable. On the scale of this plot, the uncertainties due to counting statistics are essentially zero. Points from 12 to 400 keV are from HEAO-1 scintillation counters (courtesy D. Gruber).

16.2.2 Spectrum of the background

The large area HEAO-1 detectors, in 6 months of observing, collected over 10 million counts useful for broadband spectral analysis. Figure 16.2 shows this spectrum, which is one of the best determined in all of X-ray astronomy! Between 3 and 50 keV, it has exactly the shape expected of thermal bremsstrahlung, radiation from hot gas at a temperature of one half billion degrees (40 keV). The agreement with the theoretical spectral form is excellent. There is no way a single power law, the type of spectrum emitted by nearby active galactic nuclei and quasars, will fit the data.

In the limited energy range 2–10 keV, this agreed with early observations, which were all consistent with a power law spectrum with index

≈ 0.6 . The HEAO-1 data showed that there is a ‘break’ in the spectrum at ~ 40 keV, after which the slope of the spectrum steepens considerably. The spectral shape led immediately to the hypothesis that hot gas accounts for the background emission in this energy range. No line emission was seen, but little was expected from material at so high a temperature. Also, if primordial, the gas might consist of only H and He, and there would be no lines in this energy range at any temperature.

The only place where this very hot gas might be found is the space between clusters of galaxies. A uniform but very tenuous space density of about 10^{-6} ions cm $^{-3}$ would produce the observed signal. However, the total mass involved would be enormous, more than all other known or implied mass in the Universe. This was an exciting thought.

Support for the hot-gas hypothesis came from well-measured spectra of nearby extragalactic sources. Could the background spectrum be reproduced by a collection of these sources at various red shifts? The answer was no. Clusters of galaxies are the only known thermal sources, and these have temperatures of from 2 to 8 keV – much too low to reproduce the X-ray background spectrum. The majority of extragalactic sources are active galaxies and quasars, and these have power-law spectra with average energy index 0.7 (Chapter 14). These cannot make up the background spectrum since, even when red shifted, power-law spectra are straight lines on the plot in Fig. 16.2. Well, we now know that the background is indeed due almost entirely to faint AGN.

Many are strongly absorbed by material in the host galaxy, and this absorption is the key to understanding the background spectrum. The absorbed spectrum is a power law with a break (a sudden shift to a more rapid decrease of flux with energy) at ~ 100 keV, and the spectrum of Fig. 16.2 can be explained as a sum of absorbed and unabsorbed AGN spectra over a range of red shifts. The hot-gas hypothesis was abandoned when the background was resolved into a collection of very faint point sources and the sum of their emission was shown to be about equal to the observed ‘diffuse’ signal.

16.3 Imaging the background

16.3.1 Einstein 0.3–3.5 keV, 1979

The Einstein mission carried the first sensitive X-ray telescope. Spatial resolution of $\approx 5''$ enabled the first serious search for faint sources in the diffuse background. One of the mission’s prime objectives was to ‘image the X-ray background’. The deep survey fields were certainly among the most curious to be selected for observation. The telescope was deliberately pointed at places known to be devoid of interesting, bright targets. A great deal of observing time was used looking, to the limit of the telescope’s ability, at a hitherto blank field!

The finding of individual faint sources was rather like looking at a cloud and trying to see the individual water droplets. The regions selected for study were all at high galactic latitude with low absorbing column density and contained no

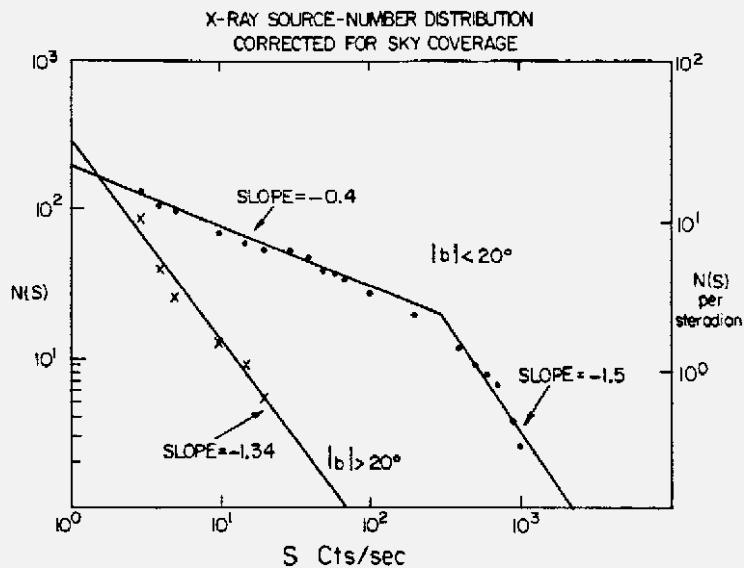
Box 16.2 | The log N-log S relation

In X-ray surveys, the number of sources increases rapidly as the flux decreases. Most sources are faint, and only the location, approximate strength and sometimes identification with an optical counterpart are known. The study of faint sources therefore proceeds using the number density and identifications. The number density, $N(S)$, is the total number of sources in a given angular range having flux greater than or equal to the value S .

Logarithms are used to give a large range on the graph.

The flux of radiation from any source falls as the inverse square of the distance. If the Universe were Euclidian and sources were distributed uniformly, the number of detectable sources would increase as the cube of the distance. In this case, the plot of $\log N$ versus $\log S$ would be a straight line with slope of $-3/2$. If all sources were

Fig 16.3 Number intensity distribution for sources found in the Uhuru survey. Sky coverage is not complete, but the 110 sources in this sample are among the brightest in the sky. $S = 1$ on the x axis corresponds to 1.7×10^{-11} ergs $\text{cm}^{-2} \text{s}^{-1}$. The energy range is 2–6 keV. The y axis scale to the right applies only to the high-galactic-latitude sources ($|b| > 20^\circ$) which follow a curve with slope $\approx -3/2$. Faint sources at low galactic latitude, if distributed uniformly in a disc, should follow a curve with slope ≈ -1 . The slope here is flatter because absorption in the galactic plane reduces the number of faint sources detected (from Matilsky et al., 1973).



distributed in a disc, as expected for those associated with the spiral arms of our Galaxy, the number of detectable sources would increase as the square of the distance. In this case the expected slope of the log N -log S plot would be -1 . The Uhuru data plotted in Fig. 16.3 contains a mixture of galactic and extragalactic sources.

Sources with high red shifts can cause the faint end of the log N -log S relation to fall below a curve with $-3/2$ slope. A clock in a system moving with respect to us at relativistic velocity appears to us to be running slow. This relativistic ‘time dilation’ causes the energy radiated by a source during one of our seconds to be less than the energy radiated during a second measured by a clock moving with the source. Thus distant, rapidly moving sources appear to be fainter than expected. Also, at high red shifts, the energy of the photons decreases considerably in travelling from source to detector. At red shift $z = 2$, for example, the recession velocity is $0.8c$, and a 3-keV detected photon would have had an energy of 9 keV when it left the source. Distant sources are thus sampled at a higher energy, and if a source spectrum falls rapidly with energy, the source will appear fainter than expected. The Chandra log N -log S curves shown in Fig. 16.4 are strongly shaped by this relativistic turnover.

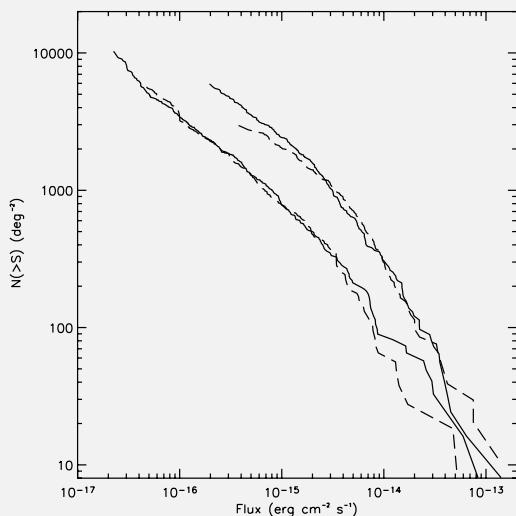


Fig 16.4 Log N -log S curves for two Chandra deep fields in two energy bands. The solid lines show source density in the north, and dashed lines, in the south. Data from the 0.5–2 keV band form the two left curves, and those at the right are for the 2–8 keV band. Source densities are the same in north and south fields, except at the faint end of the hard band, where significantly fewer sources are resolved in the south. Statistical fluctuations at the bright end of the plots are not significant (from Bauer et al., 2004).

known bright X-ray, optical or radio sources. Riccardo Giacconi and colleagues imaged several fields with both the IPC (resolution 1') and HRI (resolution 5"). Exposure times varied and ranged up to ≈ 24 hours.

The IPC was the more sensitive instrument but suffered from source confusion, a common difficulty with deep observations. If the sensitivity is high enough to detect typical sources contributing to the background, they will appear packed close together. Individual sources will not be well separated from one another. Instrumental artefacts also make analysis difficult. Small variations in the IPC detector gain put bright stripes and patches into a uniformly illuminated field. Although slight and usually negligible, these non-uniformities appeared clearly in the deep-survey exposures. Although the IPC could in theory detect faint sources, the superior spatial resolution of the HRI was necessary to confirm, separate and locate the individual sources. Using both detectors, 25 sources were clearly detected in six deep fields (Primini *et al.*, 1991). The flux limit in the energy range 1–3 keV was 5×10^{-14} ergs cm $^{-2}$ s $^{-1}$. At this level, there were 15 sources deg $^{-2}$; or, to compare with the scale of Fig. 16.3, 5×10^4 sources steradian $^{-1}$. The sources detected accounted for 16 to 25 per cent of the background flux.

Another Einstein survey, the Medium-Sensitivity Survey, was concerned with a larger fraction of the sky. In contrast with the deep surveys in which the telescope was pointed at an empty field, other observations were aimed at particular sources of interest – stars, radio sources, known quasars and so on. If the target was an X-ray source, it was positioned to appear at the centre of the field. In about half these observations, other sources also appeared within the field of view. These ‘serendipitous’ sources were completely accidental and of unknown origin. The IPC data were systematically searched, and a list of serendipitous sources at high galactic latitude was compiled (Gioia *et al.*, 1990), which included 835 previously unknown sources. Using optical spectra, 760 of these objects were identified (Stocke *et al.*, 1991). Fluxes from the extragalactic objects identified were 2–50 times greater than from the deep-survey sources and

~ 100 times less than those found in the HEAO-1 survey of Colour Plate 2.

These data on source populations had interesting implications for the X-ray background. Fifty per cent of the 51 high-latitude sources in the HEAO-1 survey were clusters, whereas only 13 per cent of those in the Medium-Sensitivity Survey were clusters. This trend continued. As fainter sources were identified, the relative number of clusters became less. X-rays from clusters of galaxies do not contribute much to the extragalactic background.

16.3.2 ROSAT, 0.2–2.4 keV, 1990

The ROSAT mirror was built for a softer energy range than the Einstein mirror. The larger angle of incidence employed made it possible to make a mirror with resolution of 5" and less scattering. Consequently, the instrumental background was less, and fainter sources were detectable. The first ROSAT observations confirmed this. Hasinger *et al.* (1993) detected sources down to a flux of 2×10^{-15} ergs cm $^{-2}$ s $^{-1}$. They found 400 sources deg $^{-2}$, implying that the slope of the log N-log S curve starts to flatten at a flux below 2×10^{-14} . This flattening was expected due to red shift and general relativity. After subtraction of obvious sources, there was evidence for structure in the background on a scale of 20', perhaps detection of structure in the early Universe. The ROSAT data brought us closer to the goal of imaging the background. The fraction of the X-ray background resolved in the 0.5–2 keV band was now 75 per cent, and as with Einstein, most of the identified sources were AGN.

16.3.3 Chandra and XMM, 0.2–8 keV, 1999

The work and dedication of Leon Van Speybroeck and a host of others produced a superb Chandra mirror with resolution of 0.5". Here was an opportunity to image sources 50 times weaker than the faintest sources seen with ROSAT. Two million seconds were spent observing the Hubble Deep Field in Ursa Major with the Chandra ACIS detector (Alexander *et al.*, 2003). The well-studied Hubble Deep Field (HDF) was chosen because optical counterparts are necessary to identify most sources. This one field, shown in Colour Plates 54

and 55, required 23 days of Chandra observing time, which, with 70 per cent efficiency, required 33 days of clock time – a big impact on the observing program! The faintest sources were detected at a level of one photon every 4 days. The flux limit was 2×10^{-17} ergs cm $^{-2}$ s $^{-1}$ in the 0.5–2 keV band. There were ≈ 9000 X-ray sources deg $^{-2}$. A deep southern field in Fornax was observed for 1 million seconds. Figure 16.4 shows the log N -log S curve for both north and south Chandra deep fields. They are almost, but not exactly, the same.

The question of how much of the background is imaged (or ‘how much of the background is due to resolved sources?’) now becomes sticky. Firstly, it is now very difficult to measure the diffuse emission remaining after subtraction of the sources. Spacecraft are in a varying (local) charged-particle environment, and there is a limit to the accuracy with which particle-induced instrumental background can be calculated. There is also a varying diffuse emission from solar-wind interactions in the Earth’s exosphere. Secondly, sources can be added and compared to the diffuse flux measured with large field-of-view detectors (Colour Plate 53), but one is soon limited by the uncertainties in these levels. There are also variations in large-scale structures which set the background level. Observations are generally summarised in two bands soft, ≈ 0.5 –2 keV; and hard, ≈ 2 –10 keV. The fraction of background now imaged by ROSAT, Chandra and XMM are calculated to be 94 per cent in the soft band and 89 per cent in the hard band (Moretti *et al.*, 2003). This includes an ≈ 2 per cent contribution from bright sources (not present in a deep pencil-beam survey) and an ≈ 5 per cent contribution from extended sources (which are often not found by point-source detection algorithms).

Again with Chandra and XMM, wide-field shallow surveys complement the deep surveys. Some source populations are too faint to appear in all-sky surveys and are sparse, so only a few are in the deep fields. Examples of such surveys are (1) the Chandra serendipitous sources (Kim *et al.*, 2004) and (2) a shallow survey of 9 deg 2 in Bootes where extensive optical data exist. One hundred twenty-five 5-ks Chandra pointings were used to study faint AGN at $z = 0.5$ –1 which comprise most of the background, particularly those associated with clusters or groups of galaxies

(Murray *et al.*, 2005); (3) The XMM Large-Scale Structure survey will cover ≈ 60 deg 2 with ≈ 500 10-ks pointings. One goal is to do a comprehensive study of cluster physics for red shifts $z < 1$. Preliminary results show a cluster density of ~ 15 deg $^{-2}$ (Pierre *et al.*, 2007).

16.3.4 Fluctuations and stacking

The deep survey fields, after removal of the obvious sources, still contain information. Fluctuations in the deep survey background can be analysed (in the same way as the HEAO-1 background fluctuations) to give information about sources that individually lie below the detection threshold. There is a granularity in the background, above the expected statistical fluctuations, implying that faint unresolved sources contribute part, perhaps all, of the isotropic background.

Fluctuation analyses have been done to extend survey data throughout the history of X-ray astronomy from Uhuru to Chandra. The level of granularity has been used to set limits on the presence of sources too faint to be seen individually. Assuming sources were distributed in intensity in a particular way, limits were placed on the source populations at levels an order of magnitude below the faintest individual sources detected.

Spectra of sources below the detection threshold can be obtained by ‘stacking’ the data. If there are e.g. 100 faint galaxies in the field, X-ray data and local background can be extracted using optical positions for each galaxy and then adding them all together. The aggregate signal can be significantly above the aggregate background and an average spectrum can be obtained. In this way, faint optical galaxies have been identified as the probable sources of the ≈ 10 per cent unresolved background in the Chandra deep fields (Worsley *et al.*, 2006).

16.4 Extragalactic source populations and cosmic variance

The type of objects found in the background depends on the flux value. Half of the 51 high-latitude sources detected in the HEAO-1 survey

were clusters. At the limit of the Einstein medium-sensitivity survey, 10^{-13} ergs cm^{-2} s^{-1} , there were 56 per cent AGN, 26 per cent stars, 12 per cent clusters and 3 per cent galaxies. In the Chandra deep fields, at a limit of 2×10^{-17} ergs cm^{-2} s^{-1} , 75 per cent of the resolved sources were AGN, 24 per cent were galaxies and 1 per cent were stars. On average, ~ 0.2 per cent should be clusters, but this amounts to approximately two per field – the numbers are small. If the faint limit were pushed to 5×10^{-18} ergs cm^{-2} s^{-1} , it is estimated that faint, star-forming galaxies would form the most numerous class of X-ray sources. At present, most of the fainter sources are AGN/quasars. Indeed, an X-ray telescope is a very efficient AGN finder.

Ten times more AGN are found in an X-ray field than in a typical optical search, so X-ray surveys have discovered many optically faint quasars. The X-ray-brighter ($> 10^{-13}$ ergs cm^{-2} s^{-1}) extragalactic sources tend to be AGN with little obscuration. The fainter sources are mostly obscured AGN and starburst galaxies. Evolution of X-ray-emitting AGN can be observed. As we look back in time by probing fainter sources at higher Z , the number of AGN per megaparsec cubed increases to a maximum at $z = 0.5–1$ and is an order of magnitude less at $z = 2–4$. However, at early times, there were a higher fraction of high-luminosity objects, which means that both number of AGN and their luminosity are time-dependent. This is an observation of the evolution of supermassive black holes and accretion processes in young galaxies. High luminosity quasars ($L_x = 10^{46}–10^{47}$ ergs s^{-1}) were formed early ($\sim 10^{10}$ years ago). Low-luminosity Seyferts ($L_x = 10^{42}–10^{43}$ ergs s^{-1}) were produced later ($\sim 5 \times 10^9$ years ago). Emission from faint galaxies increases going back to $z \approx 1$ indicating a higher rate of star formation and consequently more supernova remnants, luminous binaries and hot gas responsible for X-ray emission.

Large-scale structure causes variation in the observed distribution of faint extragalactic sources. The two Chandra deep fields, for example, show variation in the number of faint sources. In Fig. 16.4 the hard-band log N -log S curve indicates fewer faint sources in the southern field. There is also variation in spatial distribution. Figure 16.5 shows the distribution of AGN with measured red

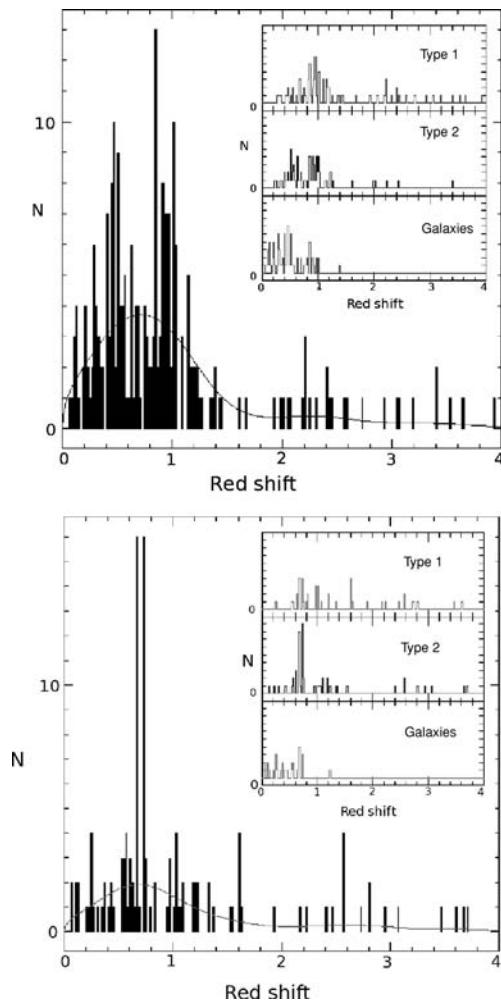


Fig 16.5 Red shift distribution for sources in the Chandra north and south deep fields. There is significant clustering in both fields at $z = 0.5–1.0$. Large plots show distribution of all sources. Small plots in boxes show a breakdown into type 1 and 2 AGN and galaxies (from Gilli et al., 2005).

shift as a function of distance. There are two clusters of sources in the north at $z = 0.5$ and 0.9 and one at $z = 0.7$ in the south. This seems to hold for both AGNs and galaxies.

16.5 | Diffuse galactic emission

16.5.1 3–30 keV, HEAO-1

After removing the obvious sources, the HEAO-1 sky map showed an excess emission from the

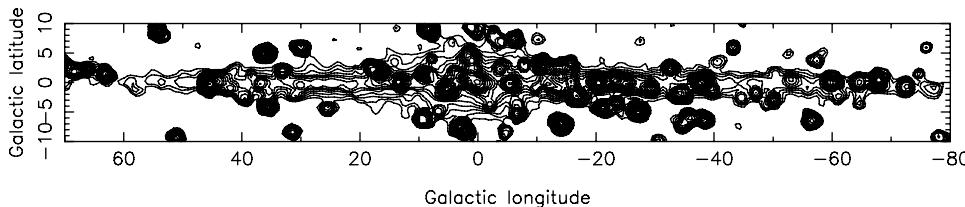


Fig 16.6 RXTE map showing the Galactic Ridge in the energy band 3–20 keV. Contour levels are logarithmic with factor of 1.4 spacing. Bright sources show the RXTE 1° resolution (from Revnivtsev *et al.*, 2006).

plane of the Galaxy. Data were analysed assuming two components: the isotropic extragalactic background and emission associated with our Galaxy. The galactic emission was bright in the galactic plane and brightest in the direction of the Galactic Centre. The relative contribution of the galactic component at the poles was only 2 per cent.

The galactic emission was fitted with a model which assumed that the radiation was from a disc-shaped region aligned with the Galactic Centre and plane. This disc contained a uniform density of emitting material. Dimensions derived, using 8 kpc as the distance to the Galactic Centre, were thickness of the disc, 5.5 kpc, and radius, 22 kpc. Thus our Galaxy, viewed from afar with an X-ray telescope, would appear as an array of ∼20 bright sources (associated with the inner spiral arms) contained in a squashed, diffuse halo with radial extent almost twice that of the region containing the bright sources.

The X-ray luminosity of this halo was ∼8 × 10³⁸ ergs s⁻¹, 30 per cent of that of the array of bright sources. The spectrum is softer than that of the extragalactic background. A thermal model with temperature of 100 million degrees (9 keV) produced a reasonable fit to the galactic-halo spectrum.

After subtraction of this galactic component, the extragalactic isotropic background remains. This result, if displayed on a sky map, has a singularly unspectacular appearance. The HEAO-1 data, binned in 9° × 9° bins, after subtraction of the galactic component, has a median standard deviation of 1.5 per cent from the mean intensity.

16.5.2 The Galactic Ridge, 2–20 keV, Ginga and RXTE

The disc component was observed extensively by Ginga (Yamauchi & Koyama, 1993) and RXTE

(Valinia & Marshall, 1998). Figure 16.6 shows an RXTE map of emission from the galactic plane within 70° of the centre. After subtraction of the ∼100 bright point sources, the resulting emission shows two components, a thin disc of width ≈0.5° and a broad halo with FWHM ≈4°. These correspond to the disc and bulge of the Galaxy. The combined spectra show a strong Fe emission line and a high-energy tail extending to ≈35 keV. Spectra were fit with a combination of a thermal model with temperature 2–3 keV and a power law with spectral index 1.8. The luminosity of the disc is ≈1 × 10³⁸ and that of the bulge ≈4 × 10³⁷ ergs s⁻¹. About 7 per cent of this energy is in the Fe line. The ‘thickness’ of the ridge/bulge calculated from RXTE data is ∼1 kpc, and the L_X is only ≈5 per cent that of the entire Milky Way.

Most of this emission can be ascribed to sources. The X-ray surface brightness follows the observed near-infrared emission, which indicates the distribution of stellar mass in the Galaxy. Cataclysmic variables, active binary stars and supernova remnants can account for most of the spectrum (Revnivtsev *et al.*, 2006). A deep Chandra pointing near the Galactic Centre confirmed this result. In the 6–7 keV band, where the Fe emission occurs, 80 per cent of the emission was resolved into discrete sources (Revnivtsev *et al.*, 2009).

At some level, some of this emission must be diffuse. Supernova remnants and star-forming regions contain large volumes of hot gas, and the remnants can supply energetic particles which interact with background photons and interstellar material in the plane. However, as is the case with the extragalactic diffuse background, new observations with more tightly collimated detectors, and consequent better identification of sources to be subtracted, have consistently led to lower limits for truly diffuse emission.

16.6 The galactic soft X-ray background, 0.2–2 keV

At lower energy, truly diffuse emission appears, but observations at such energies are much more difficult. These soft X-rays come from within the Galaxy, and there is pronounced spatial and spectral structure.

The Einstein telescope and IPC, a sensitive instrument in this energy range for sources, did not obtain useful measurements of the background. Signal-to-noise for a telescope is excellent for pointlike sources because X-rays from the source are concentrated into a small part of the detector with consequent small internal background. There is no such concentration of the diffuse background. The ROSAT mirror, however, with a 2° field of view and low-internal-background PSPC detector, was well suited for background observations. The fields of view of Chandra and XMM are so small that extensive background maps have not been attempted. The ROSAT All-Sky Survey map is currently the best soft-X-ray diffuse-emission map available. Colour Plate 56 shows the ROSAT result and the complex nature of the soft X-ray background.

16.6.1 The big picture

To appreciate some of the difficulties, let us first consider an earlier set of observations gathered by rocket-borne detectors. The observations were painstakingly conducted over a period of 8 years by William Kraushaar and colleagues, at the University of Wisconsin (McCammon *et al.*, 1983). Data from 10 rocket flights, launched from White Sands, New Mexico, and from Woomera, Australia, were used to map the 0.1–2 keV emission. Each flight scanned one-eighth of the sky with thin-window (soft X-ray) detectors collimated to a 6.5° FWHM field of view. The detectors were large-area proportional counters with thin plastic windows. The principal difficulties encountered with such observations were solar and Earth-associated backgrounds. Measurements were done at night to avoid scattered solar X-rays from the sunlit atmosphere. Windows, if thin enough to transmit soft X-rays, will also transmit low-energy electrons and

UV. These electrons can be overwhelming over large portions of a satellite orbit and are sometimes present at the top of the Earth's atmosphere.

Although the 5-min observing time of a rocket flight was certainly short compared to that of any satellite detector, the rocket had one big advantage. This was the ability to perform observations at low altitude at specific locations where the Earth's magnetic field shields the equipment from energetic charged particles. The University of Wisconsin detectors never exceeded 200 km in altitude, too low to encounter trapped particles, but sporadic fluxes of precipitating electrons were still a worry. As a precaution, ceramic magnets were built into the collimators so that low-energy electrons (<80 keV), if present, would be deflected into the walls of the collimator instead of passing through the window into the detector. The X-ray events were sorted into several energy bands by means of boron and carbon filters and electronic pulse height analysis. This also distinguished UV interactions in the detector from those of X-rays.

The ROSAT soft X-ray mapping verified and greatly improved these results with much added detail. The first 6 months of the mission were devoted to an all-sky survey. During this time the spinning spacecraft scanned the sky twice. The detector scanned a circle with radius oriented perpendicular to the Earth-Sun direction. Maximum coverage was at the ecliptic poles, and minimum coverage was along the plane of the ecliptic. The effective time pointed at each $2^\circ \times 2^\circ$ patch of sky varied from 700 s to 50 ks. This survey resulted in the detection of 50 000 sources, as shown in Colour Plate 3. After subtraction of these sources, the diffuse background remains. To ensure accuracy, a very careful, painstaking elimination of non-celestial events was necessary (Snowden *et al.*, 1995). Particle contamination and scattered solar X-rays were modelled and subtracted. Data contaminated by short-term auroral particles or solar flares were discarded. There were puzzling long-term enhancements caused by charge-exchange scattering between ions in the solar wind and hydrogen in the Earth's exosphere, the same mechanism that produces X-ray emission from comets (Chapter 4). It was difficult to know that these events were subtracted correctly.

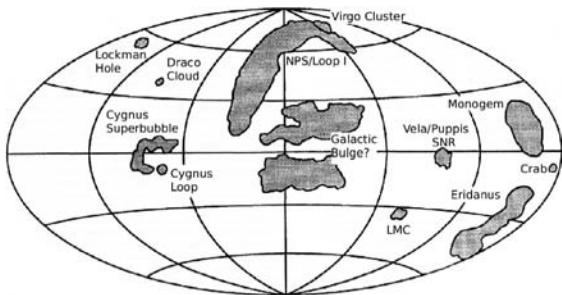


Fig 16.7 Identification of features seen in the ROSAT X-ray background maps in Colour Plate 56 (from Snowden et al., 1995).

Comparison with the rocket background maps gave confidence that contamination had indeed been removed.

Appearance of the soft X-ray sky varies greatly as the energy changes (compare Colour Plates 53 and 56). Above 2 keV, the background out of the galactic plane is isotropic and featureless, except for well-known sources. Below 2 keV, there is broad, energy-dependent structure. We discuss two energy bands, the C (or 1/4 keV) band and the M (or 3/4 keV) band. Bright pointlike sources have been removed.

16.6.2 The sky at 3/4 keV

The M (middle) bandmap, 0.45–1.2 keV, is shown at the bottom of Colour Plate 56, and Fig. 16.7 identifies some of the features. The X-ray sky is approximately isotropic with large-scale structures superposed. The map is dominated by broad emission from the central part of the Galaxy: there is an irregular feature 110° in diameter which coincides with a radio feature called Loop I. The brightest part of the rim of this loop, in the radio band, is the North Polar Spur. There is a region ~50° in diameter corresponding to the Galactic Bulge or halo. This is strongly absorbed in the galactic plane. Smaller features are the Eridanus–Orion enhancement, very faint at galactic coordinates 200° – 35°; the Cygnus Superbubble (see Chapter 7) at 80° + 5°; and various supernova remnants. These patchy features are the result of hot gas filling cavities in the ISM. The energy to heat the gas comes from supernovae and winds from early-type stars.

The 2–10 keV spectrum can be extrapolated to predict the extragalactic diffuse contribution in the M band. This contribution should be strong, with half of the flux at the poles predicted to be of extragalactic origin. This extragalactic component is completely absorbed by interstellar gas in the direction of the galactic plane. Since Colour Plate 56 shows no minimum at the galactic equator, local sources must contribute almost all of the equatorial flux. Some of this flux comes from coronal emission of dwarf G, F, K and M stars and distant, unresolved galactic superbubbles. The exact nature of all the contributing sources is not well understood.

The ROSAT PSPC, with low internal background, produced detailed maps of diffuse soft X-rays. Diffuse structures in both emission and absorption are seen on scales ranging from a few degrees to a few arcminutes. Chandra and XMM reveal thousands of faint sources in the galactic plane that contribute to what has been considered the halo emission. Separation of point source and diffuse emission is difficult. Colour Plate 38 shows part of the Chandra Galactic Centre survey and gives an appreciation of the complexity of the task.

16.6.3 The sky at 1/4 keV

At the lowest energies, the diffuse sky in the C (carbon) band is distinctly different than in the M band. This band was determined by plastic detector windows which transmit X-rays in a narrow region between ~0.15 keV and the carbon K absorption edge at 0.28 keV. Colour Plate 56 compares the rocket-measured C-band map with that of ROSAT at 1/4 keV. Note the similarity of structure in the two data sets.

The brightest regions in this band are associated roughly with the galactic poles. This is most easily interpreted as absorption of an extragalactic flux by the neutral gas in the galactic ISM. The column density of neutral hydrogen is indeed highest in the galactic plane. Although there is a pronounced anticorrelation between X-ray and 21-cm (radio emission from H atoms) sky brightness, the picture is not this simple. Although the gas density in the plane is large enough to easily absorb all the X-rays in the plane, there is appreciable emission coming from that direction. So

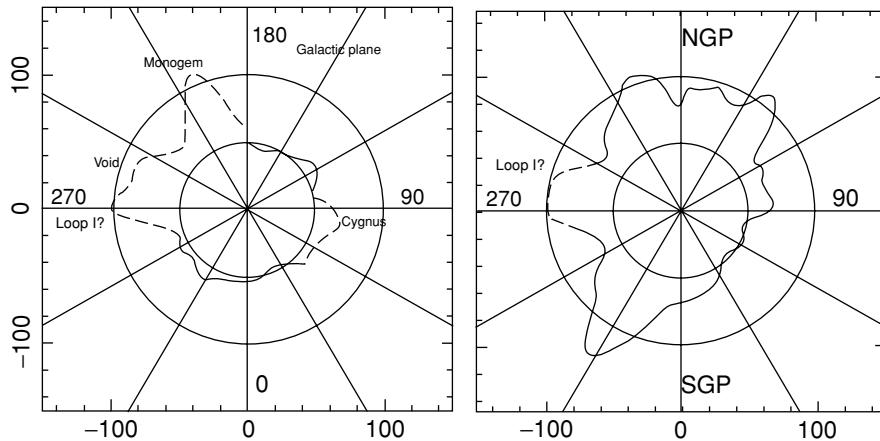


Fig 16.8 The extent of the LHB as determined by an analysis of ROSAT data. Two slices through the location of the Solar System are shown. (left) A plot of the bubble extent in the galactic plane. Inscribed numbers are galactic longitude. (right) A plot of the extent in a plane passing through the galactic poles and along longitudes 90° and 270° . The line is dashed in regions contaminated by the features indicated. The two circles have radii of 50 and 100 pc (from Snowden *et al.*, 1998).

there must also be local sources of these very soft X-rays. Also, the flux of soft X-rays in all directions is considerably higher than that calculated by an extrapolation of the smooth extragalactic background spectrum.

Where exactly is this million-degree gas which produces the soft X-ray background? Radiation from a hot galactic corona is one possibility. As discussed in Chapter 13, some other galaxies have X-ray-emitting halos which extend beyond the optical boundaries. This could well be true for our Galaxy, but even if so, the lack of complete absorption in the plane rules out this as the only source of soft X-rays. There is indeed a ‘local’ source of emission within 100 pc of the Sun. The Sun is known to be in a region where the surrounding density of neutral hydrogen is lower than usual. Apparently this is so because this volume is full of hot ionised gas. The Sun is within a bubble of hot gas in the ISM, the Local Hot Bubble (LHB). A diameter of about 100 pc and an electron density of 10^{-2} cm^{-3} would produce the necessary emission.

Shadows cast by molecular clouds reveal that some of the softest emission is nearby. The clouds are thick enough to stop most 1/4-keV X-rays but only cast partial shadows. Emission observed in front of the clouds must have a local origin. Two early examples discovered by ROSAT are the cloud

MBM 12, 90 pc distant, which barely casts a shadow whereas the Draco Cloud shown in Fig. 5.5, 300 pc distant, stops ≈ 50 per cent of the 1/4-keV emission. The Draco Cloud is at high galactic latitude and is $1/2^\circ$ by 1° in extent. The result shows that although much of the background must come from the LHB, there is an appreciable contribution from the galactic halo.

The three-dimensional extent of the LHB has been painstakingly derived from the ROSAT survey (Snowden *et al.*, 1998). Observed diffuse emission was assumed to have two sources: unabsorbed emission from the LHB and partially absorbed emission from the halo. The extent of the LHB was derived in all directions. Although irregular, it is somewhat elongated perpendicular to the galactic plane, which explains the observed anticorrelation with neutral hydrogen. As illustrated in Fig. 16.8, in a direction where the path length in hot gas is long, the path length in surrounding cool gas is small.

Spectra from the ROSAT sky survey were used to derive the temperature of gas in the LHB and the halo (Kuntz & Snowden, 2000). The LHB radiation was best characterised by a temperature of $\log T = 6.10$ ($kT = 0.11 \text{ keV}$). The halo emission is more strongly absorbed and accounts for much of the 3/4-keV background. A two-temperature model was needed to explain the spectrum of the halo.

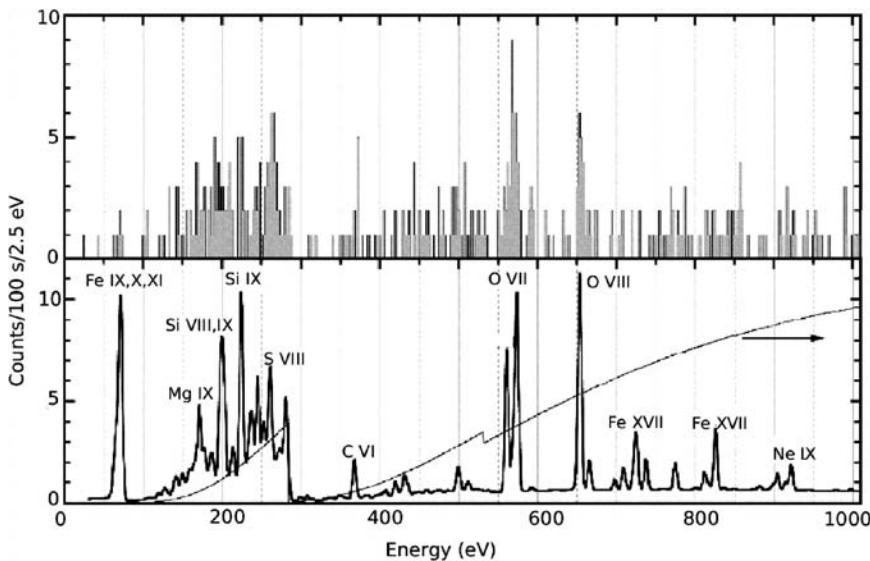


Fig 16.9 A 100-s observation of the soft X-ray background from a rocket-borne microcalorimeter. (top) Observed spectrum. (bottom) Prediction of a two-temperature thermal model with solar abundances, folded with the detector response. The solid curve shows the detector effective area. H-like and He-like oxygen ions are the dominant contribution to the 3/4-keV background and lines from Mg, Si, S and Fe to the 1/4-keV band (from McCammon *et al.*, 2002).

Best fits were obtained with $\log T = 6.06$ and 6.46 ($kT = 0.10$ and 0.25 keV).

16.6.4 Spectrum of the soft background

Emission from the LHB is thermal from gas at a temperature of about 1 million degrees. X-rays from such a gas will be almost completely in the form of spectral lines, with energies characteristic of the ions which radiate. Spectral observations of the diffuse background are consistent with this but are difficult to accomplish. Energies of the lines are low, and they are not resolved by the usual detectors (proportional counters, CCDs). The source is also distributed over a large region of the sky, so conventional dispersive spectroscopy is impossible. Emission from the LHB is also superposed on emission from the more distant halo.

A unique crystal spectrometer, the Diffuse X-ray Spectrometer (DXS), was built to detect emission lines in the range 0.15–0.28 keV (44–83 Å) from the LHB. The instrument was collimated so that at a given time, each crystal recorded only a particular energy from a particular band of the sky. The instrument was rotated so that as crystals scanned a region of sky, a range

of energy was also scanned. It was flown on a shuttle mission in January 1993 (Sanders *et al.*, 1998). The spectrum obtained indicated lines from ionised Ne, Mg, Si, S and Fe, a gas temperature of ≈ 0.11 keV and some abundances less than solar. These data were consistent with the thermal radiation expected, but emission lines were not clearly resolved.

The brighter lines were resolved by a microcalorimeter detector flown on a rocket in March 1999 (McCammon *et al.*, 2002). The field of view was ≈ 1 steradian centred at $l = 90^\circ$, $b = 60^\circ$, a region of high 1/4-keV and low 3/4-keV emission just to the left of the North Polar Spur (labelled NPS in Fig. 16.7). Figure 16.9 shows the observed spectrum and compares it with a model result which gives an idea of how data from a longer observation might appear.

A Chandra measurement of the soft diffuse background was accomplished by removing contributions from point sources and by using observations of the dark Moon to determine the detector background (Markevitch *et al.*, 2003). Data from four regions all showed a prominent blend of O VII and O VIII emission lines at ≈ 580 keV. Emission from the North Polar Spur contained lines from

Fe XVII. None of these lines were clearly resolved, but results are consistent with the microcalorimeter, DXS and ROSAT results.

16.7 | Summary

The diffuse X-ray background spans a large energy range, from 0.1 to ≥ 20 keV (Colour Plates 53 and 56). It was discovered in the 2–10 keV energy range, and here hopes of finding a new cosmological phenomenon were strongest. The uniformity of the HEAO-1 background map and the apparently thermal spectrum kept these hopes alive for some time, but high-resolution telescopes have now resolved most of the extragalactic background. There are $\sim 10^4$ faint sources deg $^{-2}$. Distant AGN and galaxies account for almost all of the extragalactic radiation.

The background below 1 keV is largely galactic in origin. There is broad spatial structure, changing rapidly with energy. There is also clearly a galactic component at higher energies, extending to at least 10 keV. The galactic component comes in part from stars and from diffuse clouds of hot gas energised by supernovae and strong stellar winds. At the lowest energy surveyed, some of the background comes from a local bubble of hot gas. There is also a fairly strong variable, soft X-ray background generated in the vicinity of Earth by the solar wind. The X-ray background is due to phenomena located over a remarkable span of distance, from the Earth's neighbourhood ($\sim 10^{-8}$ pc) to a distance of at least 5000 Mpc, where the red shift of emitted radiation is $z = 2$.

References

Observations

- Alexander, D. M., et al. (2003) The Chandra Deep Field North Survey. XIII. 2 Ms Point-Source Catalogs, *AJ* **126**, 539.
- Bauer, F. E., et al. (2004) The fall of AGN and the rise of star-forming galaxies: A close look at the Chandra Deep Field X-ray number counts, *AJ* **128**, 2048.
- Dicke, R., et al. (1965) Cosmic black-body radiation, *ApJ* **142**, 414.

- Giacconi, R., et al. (1962) Evidence for X-rays from sources outside the Solar System, *Phys Rev Lett* **9**, 439.
- Gilli, R., et al. (2005) The spatial clustering of X-ray selected AGN and galaxies in the Chandra Deep Field south and north, *A&A* **430**, 811.
- Gioia, I., et al. (1990) The Einstein Observatory extended medium survey I: X-ray data and analysis, *ApJS* **72**, 567.
- Hasinger, G., et al. (1993) A deep X-ray survey in the Lockman-Hole and the soft X-ray N-log, *A&A* **275**, 1.
- Kim, D.-W. et al. (2004) Chandra multiwavelength project. II. First results of X-ray source properties, *ApJ* **600**, 59.
- Kuntz, K. D. & S. L. Snowden (2000) Deconstructing the soft X-ray background, *ApJ* **543**, 195.
- Markevitch, M., et al. (2003) Chandra spectra of the soft X-ray diffuse background, *ApJ* **583**, 70.
- Matilsky, T., et al. (1973) The number-intensity distribution of X-ray sources observed by Uhuru, *ApJ* **181**, 753.
- McCammon, D., et al. (1983) The soft X-ray diffuse background, *ApJ* **269**, 107.
- McCammon, D., et al. (2002) A High spectral resolution observation of the soft X-ray diffuse background with thermal detectors, *ApJ* **576**, 188.
- Moretti, A., et al. (2003) The resolved fraction of the cosmic X-ray background, *ApJ* **588**, 696.
- MPE/ROSAT (1997) <http://www.xray.mpe.mpg.de/rosat/survey/sxrb/12/ass.html>.
- Murray, S., et al. (2005) XBootes: An X-ray survey of the NDWFS bootes field. I. Overview and initial results, *ApJS* **161**, 1.
- Penzias, A. & R. Wilson (1965) A measurement of excess antenna temperature at 4080 mols, *ApJ* **142**, 419.
- Pierre, M., et al. (2007) The XMM-large scale structure catalogue: X-ray sources and associated optical data. Version I, *MNRAS* **382**, 279.
- Primini, F. A., et al. (1991) The CfA Einstein Observatory extended deep X-ray survey, *ApJ* **374**, 440.
- Revnivtsev, M., et al. (2006) Origin of the galactic ridge X-ray emission, *A&A* **452**, 169.
- Revnivtsev, M., et al. (2009) Mystery of the galactic X-ray ridge solved – emission resolved into discrete sources, *Nature* **450**, 1142.
- Rothschild, R. E., et al. (1979) The cosmic X-ray experiment aboard HEAO-1, *Space Sci Instrum* **4**, 269.
- Sanders, W. T., et al. (1998) The soft X-ray background from DXS, *Lecture Notes Phys* **506**, 83.
- Snowden, S. L., et al. (1995) First maps of the soft X-ray diffuse background from the ROSAT XRT/PSCP All-Sky Survey, *ApJ* **454**, 643.
- Snowden, S. L., et al. (1997) ROSAT survey diffuse background maps II, *ApJ* **485**, 125.

- Snowden, S. L., *et al.* (1998) Progress on establishing the spatial distribution of material responsible for the 1/4 keV soft X-ray diffuse background local and halo components, *ApJ* **493**, 715.
- Stocke, J., *et al.* (1991) The Einstein Observatory extended medium survey II: The optical identifications, *ApJS* **76**, 813.
- Valinia, A. & F. E. Marshall (1998) RXTE measurement of the diffuse X-ray emission from the galactic ridge: Implications for the energetics of the interstellar medium, *ApJ* **505**, 134.
- Worsley, M. A., *et al.* (2006) Can the unresolved X-ray background be explained by the emission from the optically-detected faint galaxies of the GOODS project, *MNRAS* **368**, 1735.
- Yamauchi, S. & K. Koyama (1993) The 6.7 keV iron line distribution in the galaxy, *ApJ* **404**, 620.

Reviews

- Brandt, W. N. & G. Hasinger (2005) Deep extragalactic X-ray surveys, *ARA&Ap* **43**, 827.

Chapter 17

Gamma-ray bursts

17.1 Discovery

In 1963, to lessen the rapid proliferation of nuclear weapons, the United States and the Soviet Union signed a treaty prohibiting testing such weapons in the atmosphere and in space. To assure that there were no violations of this treaty, in the late 1960s the United States deployed a series of spacecraft, the Vela satellites, as monitors. Several spacecraft were positioned so that all of near-Earth space was always viewed by at least one set of detectors.

A nuclear explosion in space produces an intense prompt burst of X-rays, neutrons and γ rays. This signal is bright enough, and with a distinctive enough time signature, that there should be no confusion with natural events. Also, as in a supernova explosion, debris is ejected in all directions at high velocity. The primary detectors on the Vela spacecraft were designed to detect and recognise the prompt signals. Still, a clandestine test might be hidden from the prompt-burst detectors by detonating the device behind the Moon. The debris, however, which contains highly radioactive, rapidly decaying fission fragments, would be thrown from the vicinity of the explosion and free of the Moon's shadow. Gamma-ray detectors were therefore included which were capable of detecting radiation from nuclear debris.

In 1972, after 3 years of operation, the Los Alamos group responsible for the various detectors realised that the system was detecting bursts of γ rays that were real events, not some strange combination of background noise. The recording

of bursts with the same time signature in more than one satellite seemed conclusive proof that the bursts were real. Absence of signal in other on-board detectors eliminated a charged-particle origin.

Because the time of the bursts was recorded, detection by two or more widely separated spacecraft gave information concerning the location of the sources. In 1973, Ray Klebesadel, Ian Strong and Roy Olson presented their results to the astronomical community. They had seen 17 bursts in 3 years. One originated from the Sun. Neither the Earth nor the Sun was the source of the other 16. None of these was associated with known novae or supernovae. Gamma-ray energies were in the range 0.2–1.5 MeV. The shape of the bursts varied, their duration ranged from 0.1 to 30 s and they appeared to originate from different locations. No one had any idea what was causing them. This was exciting. Careful operation of detectors in a new region of wavelength and time response had resulted in an important astronomical discovery (Klebesadel *et al.*, 1973).

Box 17.1	Gamma-ray burst nomenclature
-----------------	-------------------------------------

A γ -ray burst (GRB) is named according to the date of detection. Thus the burst recorded on 21 November 2001 is called GRB 011121. Some bursts with soft spectra have been classified as X-ray flashes and are named accordingly, e.g. XRF 011030. If more than one burst should be recorded on a given day, letters are appended, e.g. GRB 960813B.

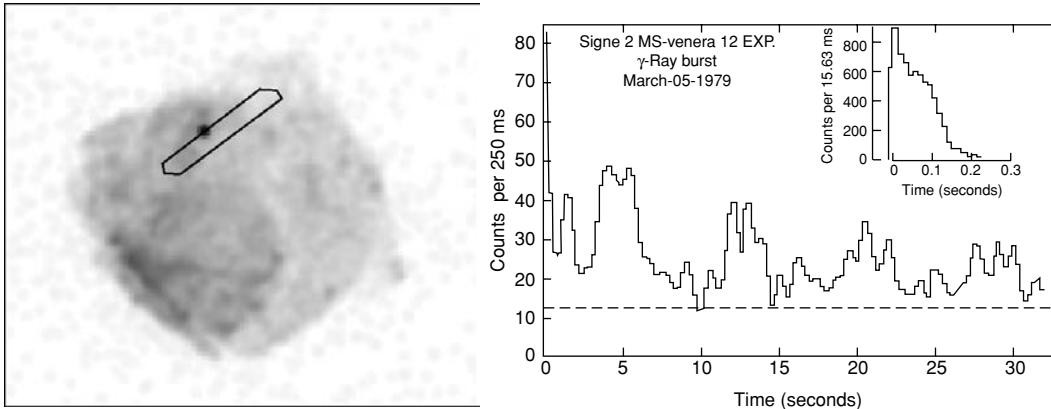


Fig 17.1 A soft γ -ray repeater. (left) The IPN location for GRB 790305 (Cline *et al.*, 1982) overlaid on a 2-hour Chandra observation of the supernova remnant N49, which is $1.4'$ in diameter. The SGR shows clearly as a bright, unresolved source within the remnant and is clearly the source of the γ -ray burst. (right) The light curve of GRB 790305 as measured by the SIGNE detector on Venera 12 in the energy band 130–205 keV. There was an intense 0.15-s burst followed by an exponentially decaying tail showing 8-s oscillations for 2 min after the burst (from Barat *et al.*, 1979).

17.2 IPN and BATSE

The next priority was to locate the origin of a burst with enough precision so that the source might be identified at other wavelengths – a star, a galaxy, perhaps something new. Accordingly, simple γ -ray detectors were piggybacked on near-Earth satellites and on probes to other planets. The resultant array of detectors with very large separations (i.e. Solar System scale) became the *Interplanetary Network*, or IPN (Atteia *et al.*, 1985). By recording the time of an event precisely, detectors on three spacecraft can locate a source at either of the two intersections of two circles drawn on the celestial sphere. Detection by four spacecraft can yield a unique location. If one satellite is close to the Earth, there is a good chance that non-detection due to occultation can help determine a unique location. Detection of a burst by two spacecraft with separation of 10^8 km and timing accuracy of 1 ms will place the source on a circle of width 10 arcmin. By 1981, ~ 100 bursts had been detected and located to an accuracy of a few arcminutes, and no obvious candidates had been found. A popular (among astrophysicists) joke at the time was that the number of different theories for origins exceeded the number of bursts detected.

17.2.1 Soft γ -ray repeaters: A digression from the main theme

On 5 March 1979, one of the brightest bursts yet seen was located by the IPN with an accuracy of 0.1 square arcmin (Cline *et al.*, 1982), and the position was within the 1.4-arcmin-diameter shell of the supernova remnant N49 in the Large Magellanic Cloud (LMC)! The initial burst lasted only 0.15 s, but weak, fading emission was observed for 2 min after the burst, and the flux was modulated with a period of 8.0 s – the telltale signature of a rotating neutron star (Fig. 17.1). In the next 4 years, 15 more bursts from this source were recorded (Golenetskii *et al.*, 1984), establishing it as a soft γ -ray repeater (SGR), the first member of a rather small class (only ~ 10 SGR are now known). These are almost certainly isolated neutron stars with ultra-strong magnetic fields similar to the anomalous X-ray pulsars described in Chapter 9, and all are in our Galaxy or the LMC. The SGR are quite distinct from most other GRBs. Their outbursts recur (i.e. they are *not* a final, catastrophic event), and the intrinsic energy release is only $\approx 10^{-7}$ that of other GRBs.

17.2.2 GRB distribution on the sky

For 25 years after discovery, the origin of GRBs remained completely unknown in spite of intense observational efforts. Many regarded the location

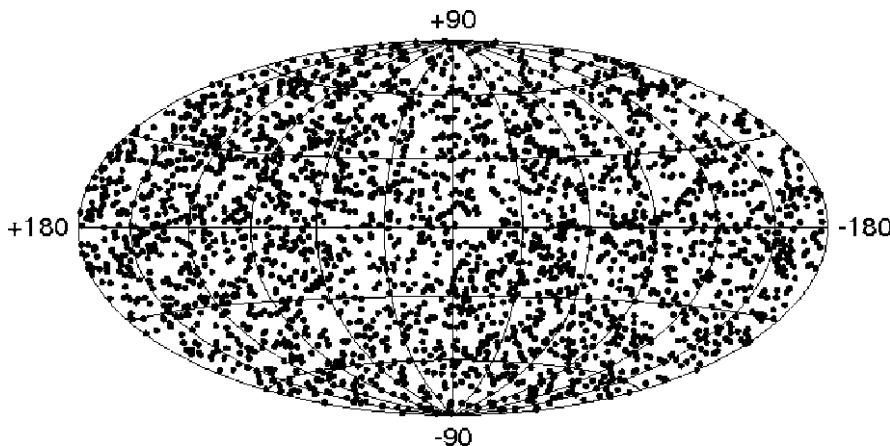


Fig 17.2 The 2704 GRBs located by the BATSE detectors. Sky coordinates are galactic latitude and longitude. The source distribution is isotropic, and there is no concentration in the galactic plane (+180 to -180 horizontal line) (NASA/MSFC/SSL, 2005a).

of GRB 790305 within N49 as an accidental superposition of a distant source and the remnant.

Lacking optical identifications, the next most important indication of origin was the distribution of sources with respect to the galactic plane. A concentration of sources close to the plane of the Milky Way would demonstrate a galactic origin. This distribution on the sky remained unclear until a sufficiently large number of GRBs had been located by Gerry Fishman and colleagues with the Burst and Transient Source Experiment (BATSE) on the Compton Gamma Ray Observatory (CGRO, the third of NASA's Great Observatories). The BATSE detectors provided locations with accuracies of $\approx 0.5^\circ$ to $\approx 10^\circ$, depending on burst strength and duration. In the 9-year life of the spacecraft, 2704 bursts were recorded, and the distribution on the sky, as illustrated in Fig. 17.2, was found to be completely uniform. The bursts were originating either in the vicinity of the Earth or at cosmological distances. But if they were at cosmological distances, then the energy required was enormous; assuming that they were radiating uniformly in all directions implied that the energy in γ rays alone could be more than $\sim 10^{53}$ ergs.

The BATSE data also revealed that although there is a great variety of structure in individual bursts, there were at least two types of events: short GRBs with duration 0.03–3 s and long GRBs with duration 7–400 s. Examples of both

are shown in Fig. 17.3 and a typical spectrum in Fig. 17.4. The number of GRBs recorded as a function of burst duration is plotted in Fig. 17.5. Although the boundary between the distributions is broad, two distinct groups are clear.

17.3 Afterglow and precise location

Detectors used to search for γ bursts, which occur at the rate of approximately one per day, have (necessarily) wide fields of view and do not measure locations accurately. To find a distant counterpart, an arcsecond location is needed to select the source of interest from the many faint objects that crowd the field. Fortunately, there is a long-lived component associated with most GRBs. This is not unexpected. Whatever the source of burst energy, it must involve material ejected at high velocity. This material will interact with the ambient medium, and the resultant heated material will radiate energy for a period of time following the burst. This is the *afterglow*.

So the key to identification of GRBs was to measure a rough, few-arcminute location quickly enough so that telescopes with much better (arcsecond) resolution could manage to image the faint, rapidly fading afterglow. Two new

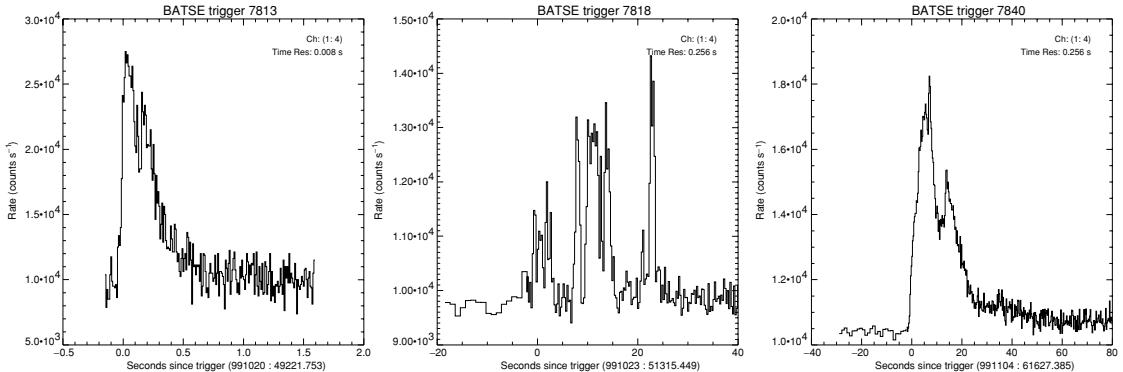


Fig 17.3 Light curves of three bursts taken almost randomly from the BATSE archive to illustrate the variety of forms observed. Timescales are linear. The left plot, spanning 2.5 s, shows a short burst lasting <0.5 s. The middle plot, spanning 60 s, shows a long burst of duration 25 s with complex structure. The right plot, spanning 120 s, shows a long burst with duration 25 s but with simple structure (NASA/MSFC/SSL, 2005b).

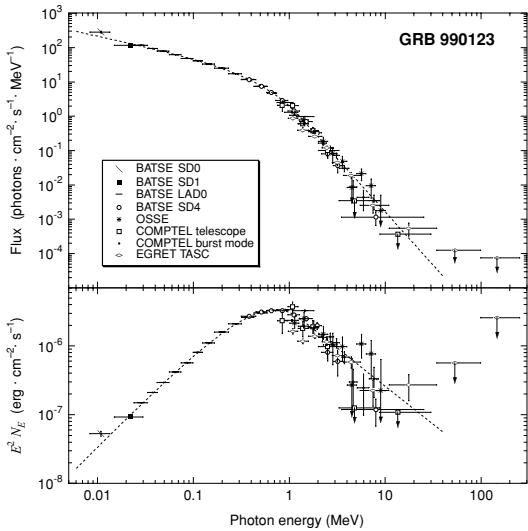


Fig 17.4 Spectrum of a bright γ -ray burst, GRB 990123. Data points are from BATSE and other GRO detectors. (top) Photon spectrum, N_E vs E , which is the usual X-ray astronomy style of presentation. (bottom) Energy flux vs energy but also multiplied by energy to condense the vertical scale for easier viewing, $E^2 N_E$ vs E (from Briggs *et al.*, 1999).

spacecraft, the Italian BeppoSAX and the US HETE 2 (Chapter 3), combined the ability to detect and locate bursts to a few arcminutes with the capacity to communicate the location within just a few minutes of the event. Rapid follow-up by other telescopes and instruments was finally possible. This led to the discovery of afterglows across a wide range of wavelengths (X-ray, optical and

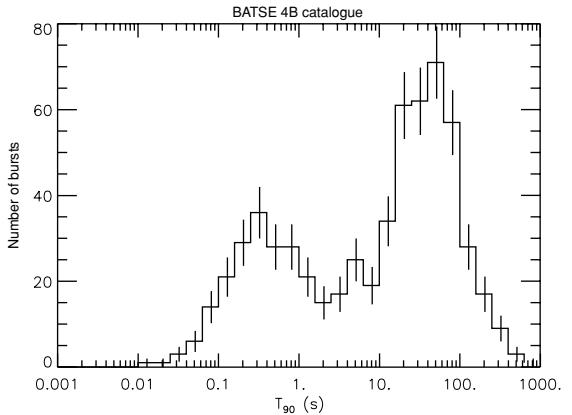


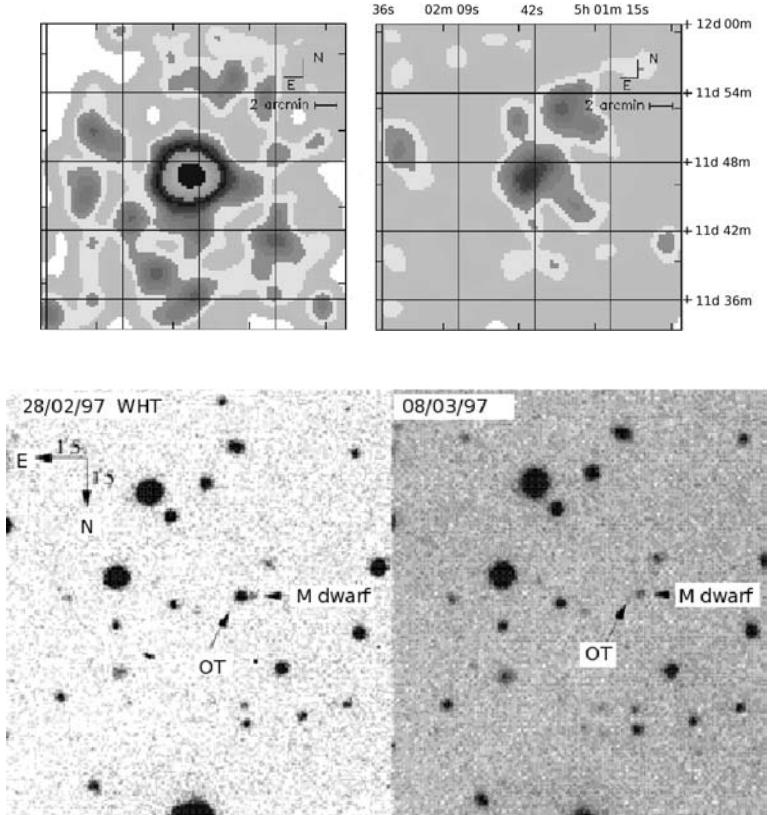
Fig 17.5 The distribution of BATSE sources as a function of burst duration from which two classes of sources are inferred. The quantity plotted, T_{90} , is the time interval from 5% to 95% of total energy radiated (NASA/MSFC/SSL, 2005c).

radio) and to the identification of host galaxies for several distant bursts. It is this afterglow that puts GRBs in a book about X-ray astronomy.

17.3.1 The first afterglow

The turning point for this field occurred on 28 February 1997, when the first measurement of a faint X-ray afterglow was made by Costa *et al.* (1997) using a BeppoSAX arcminute X-ray location of GRB 970228. This is shown in Fig. 17.6, along with the optical afterglow, which was registered by the telescopes on La Palma just 21 hours after

Fig 17.6 (top) The fading X-ray afterglow of GRB 970228 as imaged by BeppoSAX. The two observations were done 8 hours and 3 days after the burst, respectively. The field is 24' square, and the burst identification is obvious. (bottom) Pair of optical images obtained with the Isaac Newton Group of telescopes on La Palma. The WHT image shows the optical afterglow of GRB 970228 just 21 hours after the burst and is labelled 'OT' (this is the first-ever image of a GRB obtained from the ground). Four days later, there is nothing visible at that location (adapted from Costa *et al.*, 1997; van Paradijs *et al.*, 1997).



the event. This was the first time that a GRB had been seen from the ground, and van Paradijs *et al.* (1997) measured the location with arcsecond accuracy, which enabled the identification of a host galaxy. Three months later, a radio afterglow was discovered from GRB 970508 by Frail and Kulkarni *et al.* (1997) using the VLA. These were the first of many afterglow detections.

X-ray, optical and/or radio emissions have since been detected hours, days, even weeks after the initial GRB event. This radiation, although fading rapidly, often lasts long enough to permit an accurate determination of the location. Because the afterglow is faint, two observations are sometimes necessary for identification. The first, approximately a day after the burst, shows many faint sources in an area of a few square arcminutes. A week later, the second shows that one of these is a factor of 3–10 weaker, and this is the identifying signature. Colour Plate 57 illustrates the difficulty of this observation in the X-ray band. The field generally contains

many serendipitous AGN X-ray sources which are themselves variable. Finding the burst is non-trivial. In any case, arcsecond locations provided by the VLA, HST and Chandra have made possible deep optical observations which have showed conclusively that many bursts are associated with distant galaxies, some now detected with red shifts $z > 6$. These bursts were indeed at cosmological distances and may soon be recognised as the highest red shift objects known!

17.3.2 The brightest GRBs

17.3.2.1 GRB 990123

Not long after these first afterglows had galvanised the astronomical world with the revelation of the enormous distances (and luminosities) of GRBs, the GRB of 23 January 1999 produced another stunning result. This GRB was picked up by the Robotic Optical Transient Search Experiment (ROTSE), a set of 35 mm lenses attached to CCD cameras, that was operated by astronomers at the Los Alamos National Laboratory. The key to

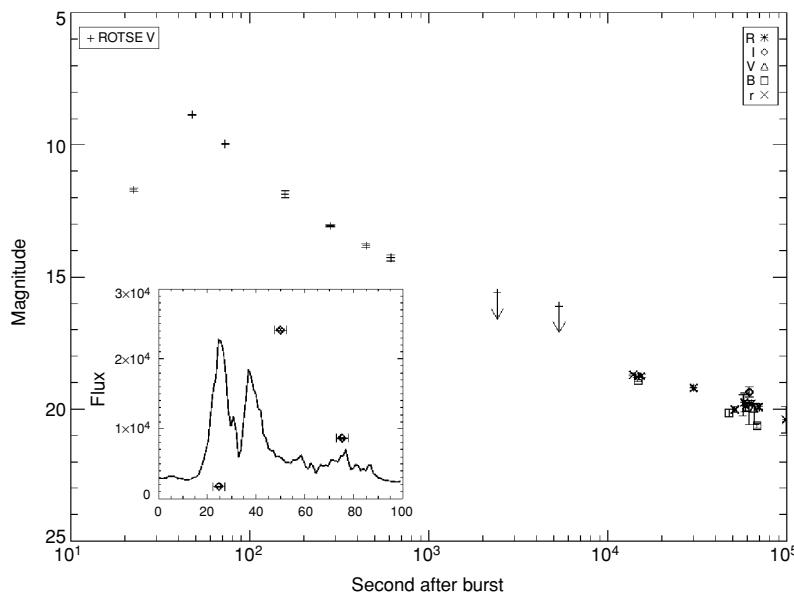


Fig 17.7 The bright optical afterglow associated with GRB 990123. This light curve was measured by the robotically operated ROTSE visible-light camera. The burst peaked at almost eighth magnitude and then faded rapidly. The inset shows how the BATSE γ -ray light curve fits in with the first three ROTSE points (from Akerlof et al., 1999).

their success was its robotic nature as it was triggered to observe the burst region by data sent from BATSE, which had detected the γ rays. As shown in Fig. 17.7, this burst was (for about 15 min) so optically bright (it peaked at eighth magnitude) that it could have been seen with a pair of binoculars or a small telescope! Nevertheless, this GRB demonstrated clearly that rapid, autonomous response by space and ground-based telescopes was to be crucial in detecting these events and learning about their true nature. Figure 17.8 is an image of the optical afterglow and also shows the faint ($z = 1.6$) host galaxy.

17.3.2.2 GRB 030329

On 29 March 2003, a huge burst was detected with strong optical and X-ray afterglow. One hour after the burst, the optical afterglow looked like a 12.4-magnitude star. This event was at $z = 0.168$, relatively nearby for a ‘cosmological’ source. The X-ray afterglow was monitored with XMM-Newton by Tiengo et al., (2004), and Fig. 17.9 shows the result. The X-ray afterglow was observable for a year, and the spectra were featureless. Because this event was so bright, we use it to illustrate the fading afterglow. The rate of fading, of course, varies from burst to burst. It is usually a power law, $\text{flux} \propto t^{-\alpha}$, where t is the time since the burst. The index, α , varies between ≈ 1 and 2 and often

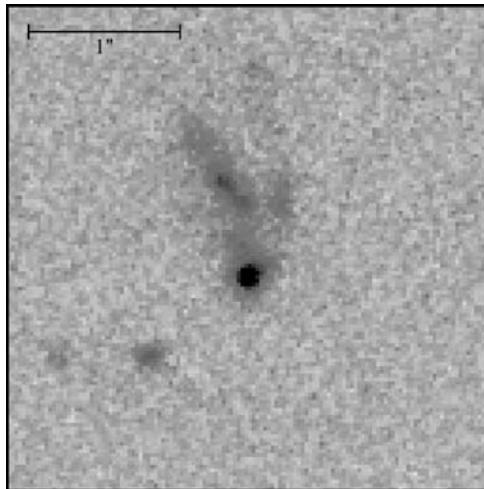


Fig 17.8 HST image of the GRB 990123 afterglow (the bright central source) and the host galaxy (from Fruchter et al., 1999).

has a value of ≈ 1.3 . Not long after the GRB (a day or so), the decay rate of most afterglow light curves increases. The time of this increase in α , the *spectral break*, is an important feature which helps determine the burst geometry, as explained in the next section.

Some afterglows are (briefly) bright enough to yield X-ray spectra. If the spectra show emission lines, red shifts can be derived from energies of the lines. Figure 17.10 shows such a spectrum

Fig 17.9 The fading afterglow of the bright burst GRB 030329 in X-ray and optical bands. Solid circles show RXTE and XMM data points. Small, closely packed circles are the optical. The break attributed to conical geometry is at 12 hours. At about 10 days, the optical emission is mostly from the 'normal' underlying supernova light curve, which is indicated by a dashed line (from Tiengo et al., 2004).

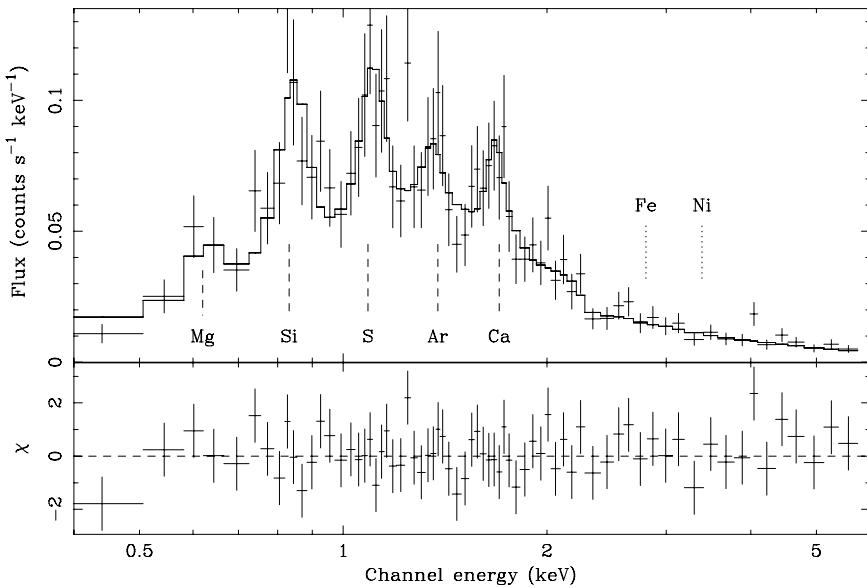
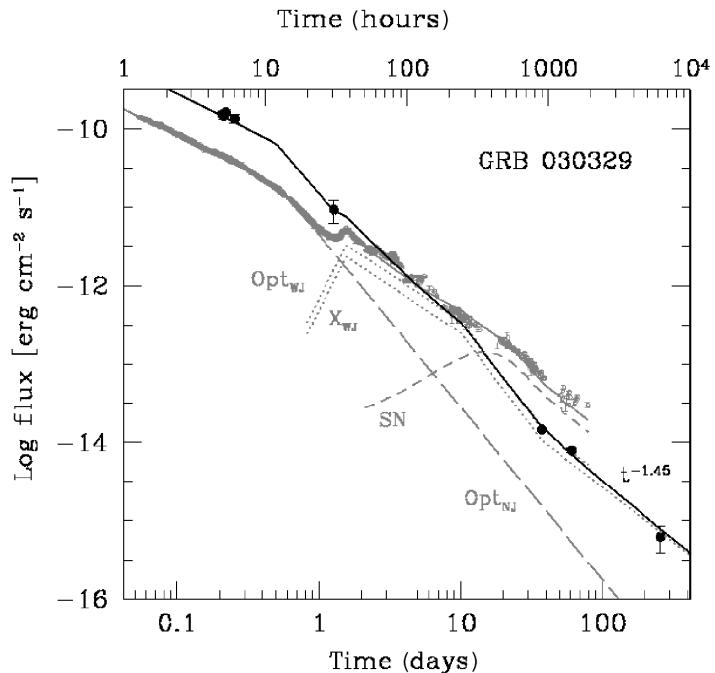


Fig 17.10 XMM Newton PN spectrum of the late afterglow of GRB 030227. Dashed lines show energies of $K\alpha$ emission lines, from the H-like ions indicated, at a red shift of 1.39 (adapted from Watson et al., 2003).

obtained with XMM by Watson et al. (2003). Line emission is apparently rare, and detection has been difficult. One needs a bright burst and must observe quickly to collect enough photons to get

a decent spectrum. Even so, most observed spectra are featureless. The spectrum of Fig. 17.10 was a 2-hour integration started 19 hours after the burst. Data taken before this time, with the same

Table 17.1 Some γ -ray bursts discussed in this chapter.

burst	fluence (ergs/cm ²)	distance	type	comment
GRB 790305	4×10^{-3}	50 kpc	SGR	SNR N49 in LMC
GRB 970228	1×10^{-5}		Long	First afterglow, Fig. 17.6
GRB 980425	4×10^{-6}	38 Mpc	Long	SN 1998bw
GRB 990123	2.7×10^{-4}	$z = 1.60$	Long	Fig. 17.8
GRB 991020	5×10^{-7}		Short	No. 7813 in Fig. 17.3
GRB 991023	3×10^{-6}		Long	No. 7818 in Fig. 17.3
GRB 991104	1.1×10^{-5}		Long	No. 7840 in Fig. 17.3
GRB 030227	7×10^{-7}	$z = 1.39$	Long	Emission lines, Fig. 17.10
GRB 030329	1×10^{-4}	$z = 0.1685$	Long	SN 2003dh
XRF 040701	4×10^{-7}			Colour Plate 57
GRB 050724	6×10^{-7}		Short	Dust halo, Colour Plate 57
XRO 080109	1×10^{-7}	27 Mpc	Breakout	SN 2008D, Fig. 8.5
GRB 080319B	6×10^{-4}	$z = 0.937$	Long	Bright optical

instrument, showed no line emission. Transient line emission, or line emission at only late times, would account for some of the null results.

17.3.2.3 GRB 080319B

This burst, caught by Swift and, at the time of the burst, actually in the field of view of two wide-field telescopes in Chile, achieved an optical magnitude of 5.3 – a burst visible to the naked eye! The γ burst lasted 50 s, and the optical event was brighter than magnitude 6 for 40 s. Light curves were measured with unprecedented accuracy and allow modelling of the prompt emission. The strong optical signal, in spite of the $z \approx 0.9$ distance of the source, was due to our fortuitous location with respect to the source geometry, as explained in the following section.

17.4 Present understanding

Apart from the SGRs mentioned earlier (and which are very different beasts from the GRBs considered in this chapter), there are two principal types of GRBs recognised in the literature: long and short GRBs, based essentially on the duration distribution of Fig. 17.5. Additionally, there are dark bursts and X-ray flashes, which may be related to the principal GRBs but are observed to have different properties for various reasons. Table 17.1 lists characteristics of γ bursts mentioned in this and other

chapters. The *fluence* listed is the time-integrated X-ray or γ -ray energy in the burst, which varies by 4 orders of magnitude for different bursts.

17.4.1 Long bursts

Most GRBs are ‘Long’, with durations of tens to hundreds of seconds. Many have been precisely located and are associated with distant galaxies with measured red shifts in the range $z = 0.5 – 6$. They are now competing with QSOs for the title of the most distant objects observable in the Cosmos (GRB 090423 has $z = 8.2!$). This is not surprising once it is recognised that GRBs can be generated from the very first massive stars (so-called ‘Population III’ stars) and that these will have very short lifetimes. Because they are so luminous, GRBs can act as probes of the high-red-shift Universe, making them extremely important astrophysical ‘illuminators’. Figure 17.11 illustrates this concept.

Two long GRBs with bright and long-lasting afterglows from relatively nearby galaxies were the key to understanding their origin. A few days after the bursts, the fading emission was observed to increase to a second maximum quite similar in shape and duration to that of classical optical supernovae. Subsequent optical spectra confirmed that part of the afterglows were indeed due to supernovae. Both produced spectra showing no H or He and weak SII and were classified as type Ic. These were GRB 980425, located at the site of SN 1998bw (Galana *et al.*, 1998), and GRB 030329,

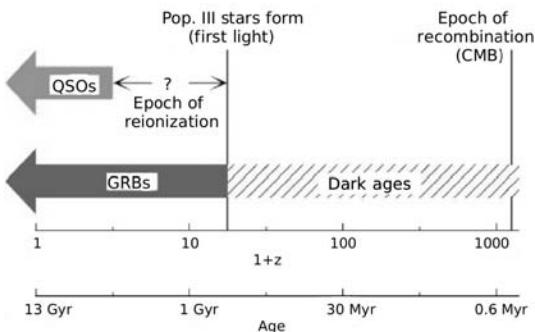


Fig I7.11 Timeline of the Universe which goes from (right) the big bang to (left) the present day 13 billion years later. Until the first stars formed, the Universe was dark. But the most massive of these first (Population III) stars rapidly produced supernovae and hence the ability to manufacture black holes via GRBs. It would take some while longer for the supermassive black holes at the centre of quasars to form (from an original by Don Lamb).

associated with SN 2003dh. The energy release of GRB 030329, however, assuming that it radiated isotropically, was calculated to be 1.4×10^{54} ergs, and it is difficult to think of a mechanism by which a supernova might release this much energy, particularly as γ rays. The answer to this difficulty requires a particular type of supernova and, especially, the effect of beaming.

Results have been compared with a model which assumes conical rather than spherical geometry (Frail *et al.*, 2001). In this model, the collapse of a massive rotating star produces a hot fireball which is ejected at relativistic velocity along the spin axis of the star. (A fireball contains a high density of particles and photons, and the photon mean-free path is short.) At first, the jet is tightly collimated with opening angle θ_j . Because the initial velocity is close to the speed of light, an interesting effect of special relativity is that we (as stationary observers) see most of the energy to be narrowly beamed along the direction of motion of the particles. At first, the beamed radiation from a particle is more tightly collimated than θ_j . As the jet interacts with the surrounding interstellar material (much of it likely ejected during the star's earlier evolutionary history), the particles decelerate, and the relativistic beam spreads. When the opening angle of the beam exceeds that of the jet, there is a change in the shape of the

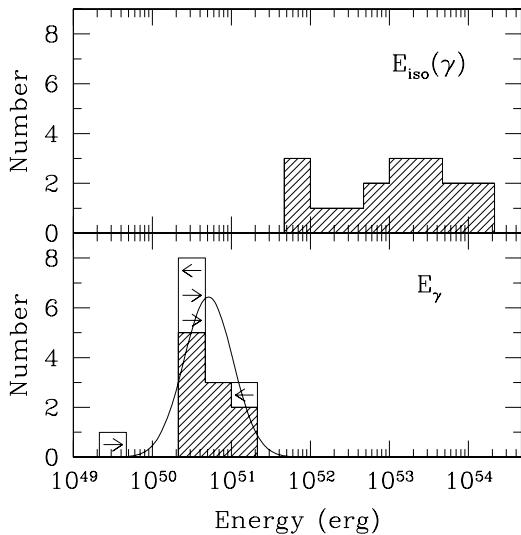


Fig I7.12 The distribution of apparent total energies of 17 GRBs with known red shifts. (top) Assumes the GRB is emitted isotropically. (bottom) Shows the energies calculated assuming the GRBs are beamed with a beam-width estimated from their observed afterglow light curves (from Frail *et al.*, 2001).

afterglow light curve. This shows as an increase in slope of the light curve called the *jet break*. This is the signature of *conical geometry*. An observer anywhere within the opening angle of the cone sees about the same radiation from the jet, whereas an observer outside this angle sees very little. Results from an analysis of 17 GRBs using this model are shown in Fig. 17.12. The break positions in their afterglow decay rates were used to calculate the jet opening angles (which varied from 1° to 25° but clustered around 4° – a jet break of ≈ 1 day) and the underlying luminosities. Although the energy releases calculated assuming isotropic emission varied from 5×10^{51} to 10^{54} ergs, those based on conical geometry were all close to 5×10^{50} ergs, quite reasonable for a 'standard' supernova.

A refinement of this concept suggests that the shape of the light curve from the bright burst GRB 080319B implies a two-jet structure. In this model, prompt events would be from material within a jet with opening angle 0.4° and afterglow from material within a surrounding, broader jet with opening angle 8° (Racusin *et al.*, 2008). The remarkable strength of the observed signal from this burst

would be due to our fortuitous location within the cone of the narrow jet.

In the supernovae above that were related to GRB events, the optical spectra indicated high intrinsic velocities and consequently very energetic explosions. It was pointed out by Paczynski (1998) that the optical afterglow of GRB 970508 was 100 times as bright as might be expected from ‘normal’ optical supernovae and that it might therefore be appropriate to call these events *hypernovae*. There is speculation that these energetic explosions might be caused by massive star core collapse to a black hole rather than to the usual neutron star.

So the long GRBs might all originate in super- or hypernovae. Gravitational collapse of a rotating star produces a black hole surrounded by an accretion disc in a short-lived analogy to the X-ray binaries of Chapter 11, the final accretion of which blows a narrow jet of relativistic material out along the spin axis. Because the beam of radiation is narrow, energy requirements of the GRB are not as severe, and the supernova explosion can supply the observed burst intensity. Since most GRBs are not aimed at us, it is estimated that the real GRB rate could be 500 times as great as that observed. This is quite reasonable given the locally estimated supernova rate of 1 per 100 years per ‘Milky Way’ – sized galaxy and that there are roughly 10^{11} such galaxies in our Universe.

17.4.2 Short GRBs

With the early IPN, several days were required to generate an accurate position of a GRB. With BeppoSAX and HETE, the response time was decreased to hours, which enabled many afterglow studies as described earlier. With the launch of Swift, a satellite specifically designed for rapid response to GRBs and other transient phenomena, the response time was reduced to just minutes. The mission was conceived in 1997 at the time the first afterglows were detected, and their rapid fading (combined with the null results in searching for afterglows from all short GRBs at that time) demonstrated the need for spacecraft follow-up to be automatic rather than involving separate missions and observatories. Consequently, Swift employs a coded-mask, wide-field γ -ray detector

which detects and locates (to an accuracy of typically 2 arcmin) about 100 GRBs per year. Having determined the GRB location, the spacecraft ‘swiftly’ slews its narrow-field X-ray and optical telescopes to study the afterglows and thereby give arcsecond positions.

Prior to Swift, no short GRBs had been located via their afterglows. By 2006, Swift had determined 18 X-ray afterglow locations from short GRBs, with a mean red shift of 0.5, which compared dramatically with the long GRB mean z of 2.3, indicating a likely different but presumably physically related origin. Indeed, with many optical afterglows now observed and host-galaxies identified, the short bursts seem to track the optical mass of galaxies, not the star-forming regions.

The short GRBs are associated with both late and early-type galaxies. They likely originate in older stellar populations and hence are not produced by some form of massive star collapse. Present thinking is that short GRBs are caused by a merger of two compact objects, perhaps two neutron stars or a neutron star and a black hole. This could produce a short-lived massive accretion disc around one of the compact objects, which would then collimate jets along the rotation axis of the binary system. Colour Plate 58 illustrates how the collision might generate magnetic fields in excess of 2×10^{15} G. These fields are one proposed mechanism for driving the relativistic GRB material in the short time before the system collapses to a black hole. A strong gravitational field is required to generate a fireball with enough energy to account for the bursts, even with conical geometry, and this is certainly potentially available in the merger of two neutron stars. Potential progenitors for short GRBs within our own Galaxy have been known for more than a quarter of a century. These are the binary radio pulsars, a handful of which are binary neutron stars. The most well known of these is B1913+16, with 7.75-hour period. The study of this binary earned Hulse and Taylor the 1993 Nobel Prize in Physics for its confirmation of the predictions of general relativity. B1913+16’s neutron stars are also two of the most accurately known stellar masses in all of astronomy (see also Chapter 1). Because of their energy loss through gravitational radiation, the orbit will decay in approximately 365 million years. The

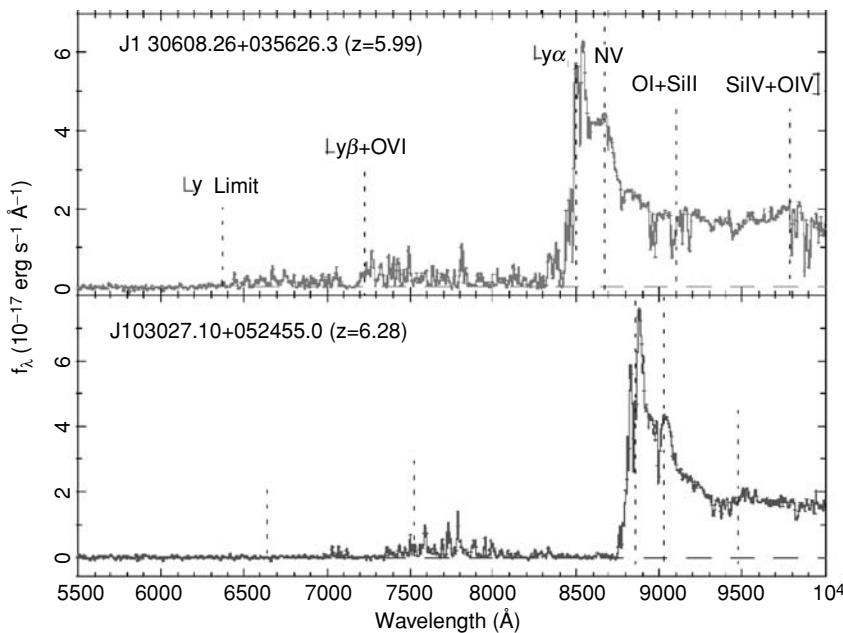


Fig 17.13 Optical spectra of high-red shift quasars as observed by the Sloan Digital Sky Survey. Note that at wavelengths shorter than that of hydrogen Ly α the continuum is almost entirely absorbed, meaning that objects with $z \geq 6$ become almost invisible in the optical and near-IR. Hence high- z GRBs might well be classified as ‘dark’ (adapted from a compilation by Don Lamb, 2005).

merger will likely create a (relatively) nearby short GRB. Another binary radio pulsar (J1756-2251) is predicted to merge in only 2 million years!

17.4.3 Dark GRBs

Some reasonably bright GRBs have no detectable optical afterglow and have consequently been classified as Dark GRBs. The cause might be an edge-on host galaxy with high extinction due to dust scattering, or perhaps the GRB is at an extreme distance ($z \geq 6$) and invisible in the optical band, as demonstrated in Fig. 17.13.

In this case, the red-shifted radiation we would observe in the optical band would have originated in the source as far-UV radiation, and the far-UV would easily be absorbed in the local ISM around the source. Such dark GRBs would offer a chance to observe the Universe at a very early time, and there has been a vigorous effort to secure precise locations with both Chandra and XMM. Recently a host galaxy has been proposed for the Dark Burst 090417 B. A dusty region with at least 15 mag of extinction seems to account for the properties of, at least, this burst.

17.4.4 X-ray flashes

Some bursts have soft spectra, peaking in the hard X-ray rather than in the γ -ray band, as shown in Fig. 17.14. These have been dubbed X-ray flashes, or XRFs. Ideas for their cause include red shift due to the high- z location of the source, a smaller Lorentz factor of the high-velocity material emitting the radiation and an off-axis view of an otherwise normal GRB jet. At present, it is not clear if XRFs are a separate class or just show one extreme of GRB spectral properties.

17.4.5 X-ray flash from SN 2008D

There has been a serendipitous observation of a brief (<5 min) X-ray flash which marked the beginning of SN 2008D in the normal spiral galaxy NGC 2770. This was observed on 9 January 2008 by Alicia Soderberg and colleagues during a Swift follow-up observation of SN 2007uy, which had erupted in the same galaxy only weeks before. The emission was caused by the rapidly expanding shock wave as it broke out of the star following the collapse of the core. This is discussed more fully in Chapter 8 and the light curve is in Fig. 8.5. We

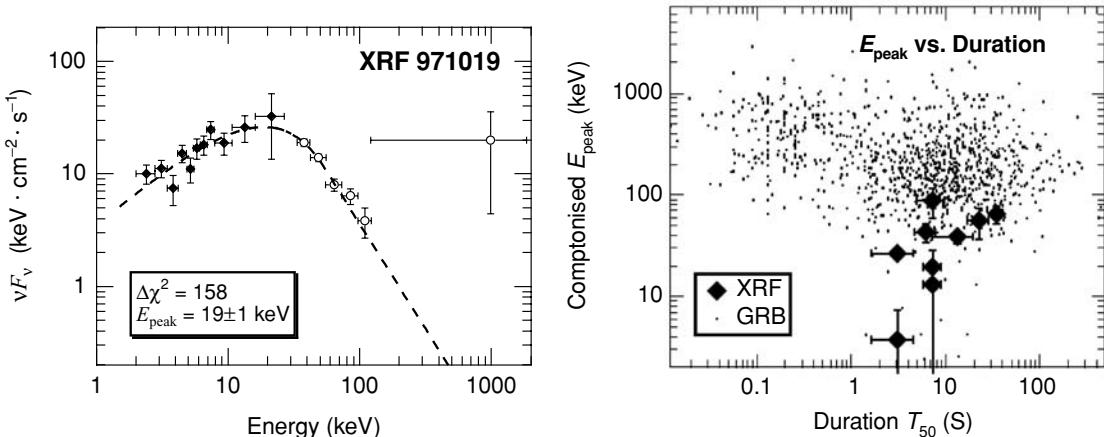


Fig 17.14 (left) Spectrum of an *X-ray flash*, so-called because of its lower energy peak (here at 20 keV) compared with GRBs. Data points are from the BeppoSAX WFC instrument and from BATSE. The presentation is the same as in Fig. 17.4, $E^2 N_E$ vs E . (right) Comparison of peak energies and durations of X-ray flashes and GRBs (adapted from Kippen et al., 2003).

mention this here and in Table 17.1 because the duration and fluence are comparable to that of a long, rather weak γ burst. The total X-ray energy release was 2×10^{46} ergs, 4 orders of magnitude below that of a γ -burst event. This is indeed a different phenomenon than the long and short γ bursts and is listed in the table as an XRO (X-ray outburst).

As a matter of interest, a short X-ray outburst was detected by one of the Vela satellites in 1969. This event also had a duration of a few minutes and occurred 2 days before the 2-month-long signal from the X-ray nova Cen XR-4 (Belian et al., 1972). Cen XR-4 was believed to be a transient X-ray binary, but no optical counterpart has been identified. The X-ray fluence was 1000 times greater than that from SN 2008D, so although possible, it is not easy to attribute this burst to a distant supernova. This mention of the Vela satellites, however, does bring this chapter full circle.

17.5 Final thoughts on GRBs

The study of GRBs has made enormous progress in the past decade. The physical constraints provided by the huge amount of energy that is observed to be emitted in such a short time interval, and at high photon energies, have pushed astrophysical

thinking to gravitational and general relativistic extremes. New observations and new theories are appearing rapidly, leading to changes in our understanding of these phenomena. Because we anticipate more surprises by the time this is published, the organisation of this chapter has been largely historical.

17.6 | The last word

Indeed, the organization of all chapters has been historical. Having been involved in this field throughout much of its history, we hope to have captured some of the excitement of the early days of X-ray astronomy. In 1968, one of us (F.D.S.) could hold in his head all known data from every X-ray source (the other, P.A.C., was still in high school!). Each observation was an exploration, with a potential gold nugget embedded within. Even though observing times were short, typically a 5-minute rocket flight, things were seen that were new and not understood. In 1972, after the launch of Uhuru, we understood that accretion onto compact objects accounted for the brightest galactic sources and that AGN and clusters of galaxies were surprisingly powerful X-ray sources. Nevertheless, one of us (P.A.C., by then a graduate student) had a little notebook with basic details of every bright

galactic X-ray source, making it easy to compare and contrast them.

In the 1980s and 1990s, EXOSAT light curves, Einstein and ROSAT images and information from many other spacecraft revealed more new phenomena. Coronal emission of stars had become a new field of study, and essentially cold bodies in the Solar System were detected in X-rays. One person, however, could no longer keep track of everything that was happening. There was an explosion of growth in X-ray information, and most of us found it a struggle to keep up with the knowledge concerning just one class of X-ray source.

X-ray observations in the new millennium are necessarily more difficult. Instruments are expensive, more complicated, and there is a long time delay between design and use of an instrument. Observations must be planned with care and approved by a committee before observing time is granted. To look to ‘see what is there’ worked very well in the 1970s but is no longer considered a valid justification for approving an observation. Furthermore, X-ray data alone are rarely sufficient to understand a source. Information at other wavelengths must be sought (sometimes it must be obtained simultaneously with the X-ray observations; at others it requires exploitation of the online Virtual Observatory) to unravel what is going on. Nevertheless, we still hope for the thrill of unexpected discoveries, and they do happen. Serendipity is a delightful feature of all aspects of astronomy.

The instruments of future missions are being designed to answer specific issues, burning questions raised largely by past observations. That answers will be found is guaranteed, and there is satisfaction in seeing them, but they are often different from what was expected. That is why this is called ‘research’. X-ray astronomers have now morphed into members of the multi-wavelength community and are not the frontier spirits they were in the past. However, X-ray astronomy has become as fundamental a part of basic astronomy tools as optical or radio observing, and we count ourselves fortunate to have been a part of the field during its maturation. And yes, we do still look forward to future observations with great enthusiasm, hoping for the thrill of the unexpected.

References

Observation

- Akerlof, C., et al. (1999) Observation of contemporaneous optical radiation from a γ -ray burst, *Nature* **398**, 400.
- Atteia, J. L., et al. (1985) A second catalog of γ -ray bursts: 1978–1980 localizations from the interplanetary network, *19th Inte Cosmic Ray Conf* **1**, 33.
- Barat, C., et al. (1979) Evidence for periodicity in a gamma ray burst, *A&A* **79**, L24.
- Belian, R. D., J. P. Conner & W. D. Evans (1972) A probable precursor to the X-ray nova Centaurus XR-4, *ApJ* **171**, L87.
- Briggs, M. S., et al. (1999) Observations of GRB 990123 by the Compton Gamma-Ray Observatory, *ApJ* **524**, 82.
- Cline, T. L., et al. (1982) Precise source location of the anomalous 1979 March 5 gamma-ray transient, *ApJ Lett* **255**, L45.
- Costa, E., et al. (1997) Discovery of the X-ray afterglow of the gamma-ray burst of February 28 1997, *Nature* **387**, 783.
- Fox, D. B. (2004) GCN GRB Observation Report #2630 http://gcn.gsfc.nasa.gov/gcn3_archive.html.
- Frail, D. & S. Kulkarni (1997) *IAU Circ.*, 6662, 1, edited by D. Green, Paris: International Astronomical Union.
- Frail, D. A., et al. (2001) Beaming in gamma-ray bursts: Evidence for a standard energy reservoir, *ApJ* **562**, L55.
- Fruchter, A. S., et al. (1999) HST and Palomar imaging of GRB 990123: Implications for the nature of gamma-ray bursts and their hosts, *ApJ* **519**, L13.
- Galama, T. J., et al. (1998) An unusual supernova in the error box of the γ -ray burst of 25 April 1998, *Nature* **395**, 670.
- Golenetskii, S., V. Ilinskii & E. Mazets (1984) Recurrent bursts in GBS0526-66, the source of the 5 March 1979 γ -ray burst, *Nature* **307**, 41.
- Kippen, R. M., et al. (2003) Spectral characteristics of X-ray flashes compared to gamma-ray bursts, in *Gamma-ray Burst and Afterglow Astronomy 2001*, 244, edited by G. Ricker & R. Vanderspele, College Park, MD: AIP.
- Klebesadel, R. W., I. B. Strong & R. A. Olson (1973) Observations of gamma-ray bursts of cosmic origin, *ApJ* **182**, L85.
- Paczynski, B. (1998) Are gamma-ray bursts in star-forming regions? *ApJ* **494**, L45.
- Price, D. J. & S. Rosswog (2006) Producing ultrastrong magnetic fields in neutron star mergers, *Science* **312**, 719.

- Racusin, J. L., et al. (2008) Broadband observations of the naked-eye γ -ray burst GRB080319B, *Nature* **455**, 183.
- Soderberg, A. M., et al. (2008) An extremely luminous X-ray outburst at the birth of a supernova, *Nature* **453**, 469.
- Tiengo, A., et al. (2004) Late evolution of the X-ray afterglow of GRB 030329, *A&A* **423**, 86.
- van Paradijs, J., et al. (1997) Transient optical emission from error box of the gamma ray burst of 28 February 1997, *Nature* **386**, 686.
- Watson, D., et al. (2003) Delayed soft X-ray emission lines in the afterglow of GRB 030227, *ApJ* **595**, L29.
- Catalogues, reviews, popular**
- Fishman, G. J. & C. A. Meegan (1995), Gamma-ray bursts, *Ann Rev Astron Ap* **33**, 415.
- Fox, D. & J. Racusin (2009) The brightest burst, *Sky Telescope* **117**, 34.
- Gehrels, N., L. Piro & P. J. T. Leonard (2002) The brightest explosions in the universe, *Sci Amer* **287**, 84.
- NASA/MSFC/SSL(2005a) <http://www.batse.msfc.nasa.gov/batse/grb/skymap/>.
- NASA/MSFC/SSL(2005b) <http://www.batse.msfc.nasa.gov/batse/grb/lightcurve/>.
- NASA/MSFC/SSL(2005c) <http://www.batse.msfc.nasa.gov/batse/grb/duration/>.

Index

- H_0 , 277, 278
1E 1207.4-5209, 132–3
3C 273, 22
3C 58, 139, cp26
3C 129, 281
3C 273, 278, 295, 297, cp49
3C 321, cp45
3C 75, cp46
40 Eridani, 66
44i Boo, 76, cp11
47 Tuc, 32, 235, 238, cp32
4U 1538-52, 59
4U 0115+63, 179, 181
4U 1323-62, 208–9
4U 1624-49, 206
4U 1626-67, 204
4U 1636-536, 214, 217–8, 228
4U 1700-37, 177, 184, 186–7
4U 1728-34, 210, 224
4U 1820-30, 218
- A0620-00 (V616 Mon), 246
A2029, 313
A2142, 315–16
accreting millisecond X-ray pulsars (AMXPs), 225
link with radio pulsars, 226
neutron star signature, 225
properties, 227
spin frequencies, 228
accretion column, 157
accretion cylinder, 183
accretion disc, 145
spectrum, 146, 148
truncation, 156
viscosity, 146
winds, 152, 153, 165
accretion disc corona (ADC) systems, 199, 207
schematic, 208
X-ray spectroscopy, 208
accretion physics, 177
accretion stream, 157
accretion-induced collapse, 167
active galactic nuclei (AGN), 277
accretion disc, 285, 297
torus, 297
activity range, 279
broadband spectrum, 297
- deep X-ray surveys, 279
fundamental plane, 301
jets, 292, 294
precessing, 296, cp46
schematic, 295, cp45
mass measurements, 291
optical spectra, 280
power spectra, 300
properties, 280
unified model, 281, cp39
winds, 301, 302–3
X-ray light-curves, 298
X-ray spectroscopy, 286
Ginga-12, 286
- active region
stars, 70–3
Sun, 63
- advection-dominated accretion flow (ADAF), 256
- AE Aqr, 146, 161, 162
- afterglow (GRB), 342–4
- age, stellar, 82
- Aldebaran, 69
- Algol, 78
- Alpha Centauri, 65
- AM CVn systems (ultra-compact binaries), 145, 164
- AM Her (=4U 1809+50), 146, 148, 157–8
X-ray light-curves, 159
X-ray spectrum, 158
- AM Her systems (polars), 155–6, 157
2-pole accretion, 158
schematic, 157
- Anomalous X-ray Pulsars (AXP), 135–6
- Antennae (colliding galaxies), 272–3
- AO 0235+16, 294
- AR Lac, 73
- Arcturus, 69
- Ark 564, 299
- asymptotic giant branch (AGB), 167
- atmosphere
absorption, 3
Titan, 47
- auroral X-rays, 40, cp7, cp8
- background spectrum, 327
- BATSE, 341
- Be stars, 186
- properties, 188
rapid rotators, 186
schematic, 188
- Beta Model, 311
- big blue bump, 298
- binary radio pulsar, 243
- binary stars, 67–74
bipolar outflow, 165
- BL Lac objects (blazars), 294, 295, 296–7
- blackbody radiation, 15–16, 211
- black hole, 242
binary schematics, cp33
Kerr, 254
- measuring masses, 244, 249
- observed properties, 244, 245
- Schwarzschild, 254
- SN explosions, 250
- spin, 254
- ultimate physics test, 243
- X-ray binaries, 242
eclipsing, 249, 265
jets, 251
power spectra, 252
quasi-periodic oscillations (QPOs), 255
spectral state transitions, 251
unified model, 253
- X-ray light-curves, 246, 247
X-ray spectra, 250, 251–2
- black hole candidates, 246
how to find, 255
- bombardment by electrons, 16
- Bondi-Hoyle accretion, 183, 184
- boundary layer, 147
- broad-line radio galaxy (BLRG), 280
- Bulge sources, 223
- Bullet Cluster, 317, cp50
- CAL 83, 165
- calorimeter, 23
- Capella, 9, 76
spectrum, cp12, cp13
- Carina Nebula, 92, cp17
- Cas A, 7, 8, 116–7, cp22
- cataclysmic variables, 143
accretion geometry, 146, 147
- chronology, 146
- Einstein X-ray survey, 146

- formation and evolution, 153–4
 observed properties, 145
 orbital period distribution, 154
 visualisation, cp28
 X-ray spectroscopy, 162
- causality, 243
 CD galaxy, 313
 Cen A (NGC5128), 293, 298, cp44
 Cen X-3, 7, 173, 183–5, 188
 Central Compact Object (CCO), 135
 central engine, 298
 Chandra Deep Field, 330–1, cp54, cp55
 Chandrasekhar Limit, 130, 167
 charge-coupled detector, 21
 chromosphere, Sun, 61
 circular polarisation, 157
 classical nova, 167
 classification of
 clusters, 307
 stars, 67
 supernova remnants, 120–1
 supernovae, 100
 X-ray binaries, 273–5
 clusters
 classification, 307
 evolution, 308
 luminosity, 315, 320–1
 mass, 320–1
 merger, 317
 relaxed, 308
 richness, 307
 temperature, 320–1
 coded mask, 24, cp5
 cold front, 315–17
 colliding galaxies, 271–3
 colliding winds, 89
 colour-colour plot, 273
 Coma Cluster, 311–13, cp47, cp48
 comets, 44–45
 common envelope, 101, 192
 Compton scattering, 17
 convection, solar, 61
 cooling flow, 314
 cooling flow spectrum, 162
 cooling rate of plasma, 15
 corona
 stellar, 69
 Sun, 61–63
 coronal hole, 63
 cosmic rays
 acceleration, 114
 rejection by detector, 20
 cosmic variance, 331–2
 cosmology, 322–3
 Crab Nebula, 6, 8, 57, 124–6, cp23
 Critical Density, 309, 322
 CTB 109, 139, cp27
 cyclotron radiation, 157
 Cyg A, 277, 292, cp42, cp43
 Cyg Loop, 8, 94, 119–20
 Cyg OB2, 85
 Cyg X-1 (HDE226868), 53, 244, 246, 247, 301
 Cyg X-2, 53, 171
 Cyg X-3, 58, 85, 188
 Dark Ages, 348
 Dark Matter, 306, 309, 322
 Delta Ori, 87, 89
 detectors, 18–24, 36
 DG Tauri, 82
 differential emission measure, 74, 76
 diffraction spectrometer, 27
 disc-instability model (DIM), 150–1
 disc-shadowing, 204
 dither, 56
 Doppler delay, 179
 Doppler imaging, 72, 76
 double degenerate binaries, 156
 double-lined spectroscopic binary, 179
 DQ Her systems, 160
 Draco Cloud, 54
 dust
 extinction, 51
 infrared emission, 91
 scattering, 51, 55–9
 dwarf novae, 143
 outbursts, 148
 X-rays, 151
 dynamo, stellar, 68
 Earth, 39–41
 eclipse imaging, 72, 73
 Eddington limit, 199
 ellipsoidal variation, 247
 elliptical galaxies, 270–2
 emission measure, 74
 equipotential surface, 175
 Eta Car, 90–2, cp16
 event horizon, 242, 286
 direct evidence, 255
 EX Hya, 143, 146, 148, 159, 160, 164, 176
 EX O0748-676, 205–6, 208, 214, 217
 EX O2030+375, 189
 expansion of Universe, 277, 278
 flare
 solar, 63–5
 stars, 77–9
 flickering, 152
 flux, 8
 focal spot, 26
 Fourier analysis, 220
 G189.6+3.3, 55
 galactic centre, 268–70
 proper motions, 290, cp41
 stellar orbits, 291
 Galactic Ridge, 333
 galactic soft background, 334–7, cp56
 Gamma-ray bursts
 afterglow, 342–4, cp57
 characteristics (table), 347
 dark, 350
 distribution, 342
 duration, 343
 energy, 348
 long, 347–9
 nomenclature, 340
 scattering, 58, cp10
 short, 349–50, cp58
 spectra, 343, 346
 gas-scintillation proportional counter, 20
 Giacconi, Riccardo, 5
 Ginga-12, 286
 GK Per, 161
 glitch, pulsar, 127–9
 globular cluster X-ray sources, 228
 distribution of orbital periods, 236
 distribution on sky, 229
 extragalactic, 239
 formation, 232, 233
 H-R diagram, 230
 locations, 232
 low L_X sources, 234
 origin and evolution, 234, 237
 properties, 231
 types of sources found, 236
 X-ray bursts, 230
 gravitational collapse, 101
 gravitational radiation (GR), 155
 gravitational red-shift, 288
 GRB 030227, 346
 GRB 030329, 345–6
 GRB 050724, 58, cp10
 GRB 090417B, 350

- GRB 790305, 341
 GRB 990123, 343–5
 GRO J1655–40, 248, 251, 252
 Groups of galaxies, 309–10
 GRS 1915+105, 250, 252, 253, 255
 GS 1826–24 (clockwork burster), 211
 burst properties, 213
 GX 13+1, 56
 GX 301–2, 185, 189–90
 GX 339–4, 251
 GX 5–1, 219–20, 223, cp9

 H2252–035, 146
 halo, scattering, 55–7
 He-like triplet, 76, 79, cp14
 Her X-1, 172–3
 Herbig Ae/Be stars, 80
 Hertzsprung–Russell diagram, 65, 67,
 80
 HESS, 253
 high-mass X-ray binaries (HMXBs),
 172, 177
 accretion geometry, 183, 189
 evolution, 191, 193
 luminosity function, 191–2, 374–5
 neutron star masses, 180, 182
 orbits, 181
 populations, 191
 properties, 191
 X-ray spectroscopy, 188, 190
 Homunculus (Eta Car), 91
 Hubble deep fields, 279
 Hubble’s law, 278
 hydrostatic equilibrium, 314

 IC 443, 55
 imaging proportional counter, 19
 imaging the background, 328
 Infrared (Eta Car), 91
 inner Lagrangian point, 148, 172, 175,
 184
 innermost stable circular orbit (ISCO),
 289, cp34
 interacting binaries, 173
 intergalactic gas, 55
 intermediate polar, 145–6, 156, 159
 spin modulation, 159
 X-ray spectra, 163
 X-ray, optical pulsations, 159
 interstellar
 dust, 51
 gas, 52
 intracluster medium (ICM), 306, 308,
 310
 inverse Compton radiation, 14, 297

 ionisation equilibrium, 112
 iron K α , 286
 distorted line profile, 288–9
 IX Vel, 152

 Jupiter, 45–47

 Kes 79, 140
 Krzeminski’s star, 7

 LMC X-1, 261–2
 LMC X-2, 262
 LMC X-3, 244–47, 255, 261
 Local Group, 259
 Local Hot Bubble, 336
 logN-logS relation, 328–9
 Lorentz gamma factor, 197
 low-ionisation nuclear emission
 region (LINER), 280
 low-mass X-ray binaries (LMXBs), 172,
 248
 accretion disc, 203
 atoll source, 219
 geometry, 204
 luminosity function, 271, 274
 mass donor, 203
 orbits, 205
 properties, 201, 203
 range of X-ray luminosities, 219
 schematic, 202
 searching for binary behaviour,
 203
 secondary star, 203
 Z source, 219
 lunar occultation, 25, 56, cp9

 M15 (NGC7078), 229, 234, 235
 AC211, 236
 M31, 259, 263–4, cp35
 M33, 259, 265, cp36
 M33 X-7, 249–50, 265
 M49, 318
 M80, 31, cp31
 M82, 266, cp37
 M83, 267–8
 M86, 318
 M87, 277, 285, 292, 293–4, 318
 Magellanic Clouds, 97, 259, 261–3
 magnetar, 135–6
 magnetic braking, 154
 magnetic caps, 161
 magnetic CVs (MCVs), 156
 magnetic field strength, 161
 properties, 156
 spin, orbital periods, 161

 magnetic field
 CV, 156
 Earth, 4
 neutron star, 123–4, 126–7
 pulsar wind nebula, 134–5
 short GRB, 349
 solar, 62
 magnetic polar caps, 184
 magnetic propellor model, 146, 161,
 162
 magnetosphere (Earth), 4
 magnetospheric radius, 156, 184
 Main Sequence, 65, 67
 Mars, 47–49
 mass function, 180, 244
 mass transfer, 175
 mass-transfer model, 150
 MCG-6-30-15, 286–7, 299, 300, 304
 microchannel plate, 20
 microquasars, 251, 281
 microwave background, 326
 Milky Way X-ray sources, 258, cp38
 Milky Way center, 268–70
 millisecond radio pulsars, 237
 Mirrors, 25
 modulation collimator, 24
 Moon
 occultation by, 25, 56, cp9
 X-rays from, 5, 10, 42–44
 moving targets, 49
 MR 2251–178, 299, 300
 MV Lyr, 144
 MXB 1735–444, 217

 N49, 341
 narrow-line radio galaxy (NLRG),
 280
 neutron star, 171
 blackbody radiation, 211
 cooling, 130–3
 general, 123, 127
 magnetic field decay, 226
 mass limit, 243
 mass-radius relation, 224
 measuring masses, 179
 structure, 130, cp24
 NGC 1399, 239
 NGC 1851, 229
 NGC 3227, 300
 NGC 3516, 300, 304
 NGC 3783, 302
 NGC 4038, 272–3
 NGC 4051, 299, 300
 NGC 4151, 277, 283, 298–9
 weighing the black hole, 284

- NGC 4258, 289, cp40
 NGC 4261, 285
 NGC 4278, 270–1
 NGC 4649, 271–2
 NGC 4697, 239
 NGC 5548, 302–3
 NGC 6440, 235
 NGC 6441, 229
 NGC 6624 (4U 1820-30), 229, 233, 237
 NGC 6712, 229
 NGC 6752, 235
 non-ejecting nova, 165
 normal galaxies (table), 260
 Nova Cyg 1975 (V1500 Cyg), 143, 157
 Nova Mus 1991, 251
 novalike, 144, 152–3
- O association, 85
 O stars, 84–89, 92
 observing moving targets, 49
 observatories, 30–36
 occultation, 25, 55–6, cp9
 OJ 287, 297
 old nova, 145
 orbit, satellite, 4
 Orion survey, 80–1, cp15
- P Cyg profile, 152
 p-dot diagram, 128
 period gap, 154
 Perseus Cluster, 319–20, cp52
 phoswich, 20
 photoelectric effect, 16
 photoionisation model, 162
 photosphere, Sun, 61
 PKS 1209-52, 132
 PKS 2121-150, 10
 planets, 39, 81
 Polar, 145–6
 polarization, 23
 pre-main sequence stars, 79–81
 proportional counter, 18
 Proxima Centauri, 65, 77
 PSR B1913+16, 181, 349
 PSR J1747-2958, 134
 pulsar spin-down, 126–7
 pulsar wind nebula, 133–5
- QQ Vul, 159
 quasar (QSO), 277
 broad absorption line (BAL), 301
 images, 277, 279
- quasi-periodic oscillations (QPOs), 153, 219
 beat-frequency model, 222
 kilohertz QPOs, 223
 twin-peaks, 225
 link to black holes, 225
 power spectra, 223
- Rapid Burster, 215
 burst properties, 216
 magnetospheric gate model, 216
- rapid variability, 29, 152
 Rayleigh-Taylor instability, 109–10
 relativistic beaming, 288
 relaxation oscillator, 216
 reverberation, 283
 Roche geometry, 247
 Roche lobe, 172, 184
 Roche surface (lobe), 148, 154
 Röentgen, Wilhelm, 1, cp4
 RS CVn, 70–4
 RX J185635-3754, 132
 RX J0513.9-6951
 light-curve, 166–7
 X-ray spectrum, 168
 RX J0527.8-6954, 165
 RX J0806+15, 146, 164
- S-curve, 150
 S5 0716+714:blazar light-curve, 296
 Saturn, 47
 SAX J1808.4-3658, 226
 scintillation counter, 20
 Sco X-1, 5–6, 171, 203
 optical spectrum, 172
 secondary (companion) star, 148, 247
 Sedov solution, 107
 Seyfert galaxy, 277–8, 280
 types 1 and 2, 283
 Sgr A*, 269
 shortest orbital period, 233
 slat collimator, 24
 Small Magellanic Cloud (SMC), 191
 SMC X-1, 179, 181
 SN 1006, 115, cp21
 SN 1972E, 98
 SN 1987A, 97, 103–6
 SN 1993J, 105–7
 SN 2008D, 103, 350
 SNR 0102-72.3, 116–8
 SNR 0540-69.3, 139, cp25
 Soft Gamma-ray Repeater (SGR), 135–6, 341
 soft X-ray puzzle, 157
- soft X-ray transients (SXTs), 246
 solar wind, 4, 61, 86
 solid state detector, 21
 source population, 273
 South Atlantic Anomaly, 4
 spacecraft
 ASCA, 32
 BeppoSAX, 33
 Chandra, 34, cp6
 Einstein (HEAO B), 31
 EXOSAT, 30
 Granat, 31
 HEAO 1, 30
 INTEGRAL, 33, cp5
 IPN, 341
 ROSAT, 31
 Rossi XTE, 32
 SAS 3, 8
 Suzaku, 36
 Swift, 35
 Uhuru, 8, 30
 XMM Newton, 35
 spectral resolution, 27
 SS Cyg, 8, 143–4, 146, 151–3
 SS 433, 30, 192, 281
 discovery, 194
 geometry, 200
 kinematic model, 197
 moving lines, 195–6
 relativistic jets, 196
 VLBA image, 198
 X-ray binary, 198
 X-ray image, 198
 X-ray spectroscopy, 199, 201
 starburst galaxies, 266–7
 starburst galaxies (table), 260
 stars
 age, 82
 jets, 81–2
 pre Main Sequence, 66, 79–81
 rotation, 81
 spectra, 74–76, 88
 spots, 70–1
 variable, 69
 stellar wind, 86–7, 174, 182, 184
 SU UMa system
 superoutburst, 144
 VW Hyi, 144
 Sun, 60–65
 sunspots, 62
 Sunyaev-Zeldovitch (SZ) effect, 322
 super-Eddington, 146
 superbubbles, 93–95
 superluminal expansion, 253

- supermassive black hole, 283
 binary, 297
- supernova, 143
 classification of, 100
 evolution of star to, 99
 historical, 98
 shock breakout, 103
 Type I, 99, 100–1
 Type Ia, 101, 167
 Type II, 99, 101–2
- supernova remnants
 blast wave, 107
 ejecta, 107–11
 evolution, 106–11
 M33, 265
 Magellanic Clouds, 139, 262
 reverse shock, 107
- supersoft source (SSS), 145–6, 164, 275
 X-ray spectra, 165
- survey
 Ariel 5, 9
 Einstein, 65
 HEAO 1, 9, 326–7, cp2, cp53
 rocket, 8
 ROSAT, 10, 330, cp3, cp56
 symbiotic binaries, 145, 167
 synchrotron radiation, 13–4, 134, 253, 297
- T Tauri stars, 80, 81
- telescope
 effective area, 27
 image quality, 26
 mirrors, 25–6
 Wolter, 26
- Terrier-Sandhawk rocket, 9, cp1
- thermal X-rays, 12–13, 14, 15
- Thomson cross-section, 199
- time dilation, 197
- timing, 29
- Titan, 47
- transverse Doppler shift, 197, 288
- Trumpler 14, 92, cp18
 TV Col (2A0526-328), 159
 TW Hya, 81, cp14
 Tycho's SNR, 111–15, cp19, cp20
- U Gem, 143–4, 146, 151, 153
- Ultra Luminous XRB, 275
- unstable mass transfer, 165
- UV Ceti, 77
- UV delay, 151
- UX Eridani, 73
- UX UMa, 145, 153
- V1974 Cyg, 167
- V2051 Oph, 147
- V404 Cyg (GS2023+338), 246
 radial velocity curve, 249
- Van Speybroeck, Leon, 34
- variability timescales, 29, 222
- variable stars, 69
- Vega, 69
- Veil Nebula, 119
- Vela Pulsar, 136–9
- Vela Satellites, 340
- Vela SNR, 8, 136–9
- Vela X-1, 29, 177, 185, 188, 190
- velocity dispersion
 stars in galaxy centres, 292
 galaxies in clusters, 307
- Venus, 47–49
- Virgo Cluster, 317–19, cp51
- virial radius, 309
- Virial Theorem, 309
- VW Hya, 144
- VY Scl system
 anti-dwarf nova, 144
 MV Lyr, 144
- W50 (SNR), 194
- water masers, 289
- white dwarf, 100–1, 143
- Wolf-Rayet stars, 92, 192–3, 250
- Wolter telescope, 26
- WR 140, 89, 90
- X-ray binaries, 171
 discovery of binary behaviour, 171
 galactic distribution, 177, 178
 properties, 176
 schematic, 174
- X-ray bursters, 209
 burst oscillations, 227
 dynamic power spectra, 228
 distribution of peak fluxes, 213
 optical bursts, 217
 photospheric radius expansion
 (PRE) bursts, 213
- spectroscopy, 214
- standard candle, 211
- superbursts, 218
- thermonuclear flash model, 210–11
 schematic, 212
- X-ray dippers, 204, 209
- X-ray Flash, 350–1
- X-ray fluorescence, 287
- X-ray novae (XRN), 246
- X-ray pulsars, 7, 177
 Corbet diagram, 189
 magnetosphere, 184
 orbital parameters, 179
 spinning up and down, 182
- X-rays
 attenuation, 16
 detectors, 18–24, 36
 discovery, 1, 5
 properties, 2
 reflection, 25
 scattering, 42
 shadows, 53
 transmission through ISM, 52–4
- X1822-371, 206–8
- Z Cam system, standstill, 144
- Zeta Pup, 87, 89