

Stars, Galaxies, and Beyond

Summary of notes and materials related to University of Washington astronomy courses:

ASTR 322 The Contents of Our Galaxy (Winter 2012, Professor Paula Szkody=PXS) &

ASTR 323 Extragalactic Astronomy And Cosmology (Spring 2012, Professor Željko Ivezić=ZXI).

Summary by Michael C. McGoodwin=MCM.

Content last updated 6/29/2012



Rotated image of the Whirlpool Galaxy M51 (NGC 5194)¹ from Hubble Space Telescope HST, with Companion Galaxy NGC 5195 (upper left), located in constellation Canes Venatici, January 2005. Galaxy is at 9.6 Megaparsec (Mpc)= 31.3×10^6 ly, width 9.6 arcmin, area ~27 square kiloparsecs (kpc²)

¹ NGC = New General Catalog, http://en.wikipedia.org/wiki/New_General_Catalogue

² <http://hubblesite.org/newscenter/archive/releases/2005/12/image/a/>

Table of Contents

Introduction	3
Useful Symbols, Abbreviations and Web Links.....	4
Basic Physical Quantities for the Sun and the Earth	6
Basic Astronomical Terms, Concepts, and Tools (Chapter 1)	9
Distance Measures	9
Time Measures	10
Sky Position, Coordinate System, Motion and Rotation, Ecliptic, and Sidereal Terms	12
Celestial Mechanics (Chapter 2).....	15
Kepler's Original Laws of Planetary Motion.....	15
Laws of Gravity and Motion.....	16
Light Properties (Pre-QM and Early QM); Distance and Magnitudes (Chapter 3)	18
Astronomical Distance, Luminosity, Magnitude.....	18
Magnitude and Distance Modulus.....	19
Light (Photon) Wave vs. Particle Properties	20
Color and Color Index	25
The Special Theory of Relativity (Chapter 4, omitted)	28
Quantum Properties of Light and Matter; Spectroscopy (Chapter 5)	28
Spectroscopy	28
Light Interactions and Atomic Models	29
Quantum Mechanical Considerations	31
Telescopes (Chapter 6).....	33
Binary Systems and Stellar Parameters (Chapter 7)	35
Mass Determination Using Visual Binaries.....	35
Eclipsing Spectroscopic Binaries.....	36
The Classification of Stellar Spectra (Chapter 8)	39
Harvard Classification.....	39
Statistical Mechanical Considerations	40
The Hertzsprung-Russell (H-R or HD) Diagram	44
Stellar Atmospheres (Chapter 9, omitted).....	49
The Interiors of Stars (Chapter 10).....	50
Basic Equations.....	50
Energy Transport and Thermodynamics.....	55
The Main Sequence.....	57
The Sun (Chapter 11)	59
The Solar Interior.....	59
The Solar Atmosphere.....	61
Miscellaneous Solar Topics	66
Solar Cycle and Miscellaneous Variable Solar Activity	67
The Interstellar Medium; Protostar and Early Star Formation (Chapter 12)	70
Interstellar Dust and Gas.....	70
Protostars	73
Pre-Main-Sequence Evolution, Young Stellar Objects, and the ZAMS	78
Main Sequence and Post-Main-Sequence Stellar Evolution (Chapter 13)	86
Evolution of the Main Sequence	86
Late Stages of Stellar Evolution (Post-Main Sequence).....	90
Stellar Clusters.....	94
Stellar Pulsation (Chapter 14).....	97
Key Observations of Pulsating Stars.....	97
The Physics of Stellar Pulsation	104
Modeling Stellar Pulsation	106
Nonradial Stellar Pulsation	107
Helioseismology and Asteroseismology	109
The Fate of Massive Stars (Chapter 15).....	110
Post-Main Sequence Massive Stars That Evolve To Supernovae.....	110
Supernovae (Supernovas).....	115
Gamma-Ray Bursts GRBs and Associated X-ray Emissions	122
Cosmic Rays and Solar Energetic Particles	126
The Degenerate Remnants of Stars (Chapter 16)	129
White Dwarfs (WDs)	129
The Physics of Degenerate Matter	131

Neutron Stars (NS).....	135
Pulsars	138
General Relativity and Black Holes (Chapter 17, mostly omitted)	147
Close Binary Star Systems (Chapter 18)	148
Gravity Potentials in a Close Binary Star System.....	148
Circumstellar Accretion Disks	151
A Survey of Interacting Binary Systems.....	153
White Dwarfs in Semidetached Binaries	156
Type Ia Supernovae	161
Neutron Stars and Black Holes in Binaries	161
The Solar System & Planetary Systems (Chapters 19 through 23, omitted)	173
The Milky Way Galaxy (Chapter 24)	173
Milky Way Fundamentals, Morphology, and Coordinates.....	176
Kinematics of the Milky Way	183
Galactic Center	187
The Nature of Galaxies (Chapter 25)	189
Spiral and Irregular Galaxies	190
Elliptical and SO Galaxies.....	197
Relative Numbers of Galaxies by Hubble Type	200
Galactic Evolution (Chapter 26)	203
Interaction of Galaxies	203
The Formation of Galaxies	207
The Structure of the Universe (Chapter 27).....	209
Summary of Distance Measuring Methods and The Extragalactic Distance Scale	209
The Expansion of the Universe	215
Groups and Clusters of Galaxies	218
Active Galaxies (Chapter 28).....	226
Classification and Characteristics of Active Galaxies	226
A Unified Model of Active Galactic Nuclei (including Quasars)	237
Radio Lobes and Jets	243
Using Quasars to Probe the Universe	244
Cosmology (Chapter 29, omitted)	249
The Early Universe (Chapter 30, omitted)	249

Introduction

I have prepared this summary to assist in learning some of the materials relevant to the courses named. These 300-level (junior) college courses have provided a satisfying opportunity to take a more detailed but still manageable look at astronomy and astrophysics applying to space well beyond the solar system. It is likely that humankind will not reach other stars or galaxies in the foreseeable future, but our understanding of our place in the universe is remarkably enriched by pursuing this knowledge. I thank professors Szkody and Ivezić³ for allowing me this opportunity to explore these many important topics in astrophysics.

Sources: The materials in this summary derive from the lecture notes, the assigned textbook, and many Web and published scientific articles and other sources as noted. I do not consider Wikipedia to be a definitive authority on any subject, but it is often quite useful as a starting point for more authoritative exploration, and I have included a number of citations from it. In some cases, I have added emphasis or needed punctuation to quotations. Diagrams used here are not directly copied from the textbook or from our lecture notes, but have been obtained from the quoted primary or other Web sources. I have in most cases omitted derivations of formulas and certain other relevant details such as most error ranges, but have tried to present the most important conclusions in compact form, and always where to go to read the details. For images and diagrams, I have where possible selected color examples, for clarity, didactic impact, and esthetics, but have used monochrome sources where superior in content. Images are often shrunk but are still presented in full resolution, so they can be viewed better by zooming in where desired.

³ English search terms: Zeljko Ivezic [ZXI]

Textbooks and Key Resources: The main textbook used in these courses in 2012 is *An Introduction to Modern Astrophysics*, 2nd Edition, by Bradley W. Carroll, Dale A. Ostlie, 2007 (my copy is 6th printing 2011), Pearson/Addison Wesley, hereafter abbreviated *IMA2*. References herein to chapters refer to that textbook, but a great deal of information I have included under these chapter headings is from the lectures and especially from other sources, so may not have been stated as such in the textbook (nor have I always kept the order in which material is presented in the textbook). This outline summary should not be used in place of buying this excellent textbook (which is purported to be for sophomore level astrophysics). The textbook covers many more topics, is vastly more detailed in what it covers, and is well worth the high price. A 3rd edition would be eagerly received and is needed to incorporate some of the exploding new astronomical knowledge.

I also refer occasionally to Burke BF and Graham-Smith F, *An Introduction to Radio Astronomy*, 3rd Ed., Cambridge, 2010, hereafter abbreviated *IRA3*.

DOI (Digital Object Identifier) citations may be resolved if necessary at <http://dx.doi.org/>. Enter the DOI without the “DOI:”. For instance, the article with “DOI: 10.1146/annurev-astro-082708-101642” may be found at <http://dx.doi.org/10.1146/annurev-astro-082708-101642>.

Etymology and definitions in part derive from the *Oxford English Dictionary*, online version accessed online January–June 2012, hereafter abbreviated *OED*.

I have generally favored MKS SI units (as does the textbook) rather than the CGS system of units preferred by many astronomers.

Author: I am a retired physician and an auditing student, and claim no expertise in this field. I have included material cited from many scholarly and technical articles, but I do not pretend to have mastered the contents of these articles (though I enjoy looking through them, trying to learn what I can).

Copyrights: You may use this document for educational non-commercial purposes. Please be aware that it contains some copyrighted material. If you incorporate this document or parts of it into a presentation or website etc., please acknowledge my authorship of the document, the document’s URL

http://www.mcgoodwin.net/pages/astrophysics_2012.pdf,

as well as the external sources I have cited. This is a not-for-profit informal personal study aid—if you publish any of it you should consider securing permission to publish the materials from other authors that I have included. If you are an author or copyright holder and object to anything I have included as beyond “fair use”, please advise me and I will make appropriate corrections.

This file will not be updated beyond 2012, so gradual obsolescence is to be expected.

Constructive corrections and clarifications would always be appreciated. Send these to:
MCM at McGoodwin period NET (please convert to standard format when using)

Useful Symbols, Abbreviations and Web Links

Characters and symbols useful in writing this document

Greek 12: $\Alpha \alpha \Beta \beta \Gamma \gamma \Delta \delta \Epsilon \epsilon \Zeta \zeta \Eta \eta \Theta \theta \Iota \iota \Kappa \kappa \Lambda \lambda \Mu \mu \Nu \nu \Xi \xi \O \o \Pi \pi \Rho \rho \Sigma \sigma \Tau \tau \Phi \phi \Chi \chi \Psi \psi \Omega \omega$

Greek 10: $\Alpha \alpha \Beta \beta \Gamma \gamma \Delta \delta \Epsilon \epsilon \Zeta \zeta \Eta \eta \Theta \theta \Iota \iota \Kappa \kappa \Lambda \lambda \Mu \mu \Nu \nu \Xi \xi \O \o \Pi \pi \Rho \rho \Sigma \sigma \Tau \tau \Phi \phi \Chi \chi \Psi \psi \Omega \omega$

Math & Misc. 12: $\circ \cong \approx \neq \equiv \partial \sqrt{\bullet} \int \pm \P \leq \geq \Rightarrow \angle \propto \nabla * \times ^\wedge \perp \parallel \odot \oplus ' \diamond ' \langle \rangle \hbar \infty *$

Math & Misc. 10: $\circ \cong \approx \neq \equiv \partial \sqrt{\bullet} \int \pm \P \leq \geq \Rightarrow \angle \propto \nabla * \times ^\wedge \perp \parallel \odot \oplus ' \diamond ' \langle \rangle \hbar \infty *$

Abbreviations, Acronyms, and Definitions

Most abbreviations and acronyms are spelled out at first use, and definitions are similarly worked into the text at the first or most substantive use of a term.

Astronomy Links from Professors Szkody and Ivezić, and Other Sources

These are links from Professor Szkody, Professor Ivezić, and otherwise. There are also many links on specific topics given throughout the document.

♦ = Webpages I found to be of greater interest; ♦♦ = the very best or most useful to me

- [All-Sky Milky Way Panorama](#) (Axel Mellinger, can vary tilt)
- [American Astronomical Society AAS](#)
- [American Association of Variable Star Observers AAVSO](#)
- [Cepheid Variable Stars in LMC Site](#) (see also [here](#))
- [Chandra X-ray Observatory \(CXO, 1999–\)](#): one of NASA's 3 remaining Great Observatories in space
- [Constellations](#)
- [Cosmic Calculator](#) (converts redshifts to comoving distance, light travel times, etc.)
- [Eclipsing Binary Stars Luminosity Simulation](#) (needs JAVA)
- [Elements Optical Absorption/Emission Lines](#) (needs JAVA)
- [Galactic Center](#)
- [Galactic-Scale Interactions](#) (MPEG movies)
- [General Catalog of Variable Stars GCVS](#)
- [Giant Magellan Telescope GMT](#) (future ~2020, seven 8.4 m segments, Las Campanas Obs., Chile)
- [Guide Star Catalog GSC](#) (14th mag)
- <http://heasarc.gsfc.nasa.gov/> High Energy Astrophysics Science Archive Research Center HEASARC (NASA missions)
- [Hipparcos Space Astrometry Mission](#) (ESA, high precision star positions datasets)
- [James Webb Space Telescope JWST](#) (future NASA mission ~2018, IR-optimized 6 m)
- [Large Synoptic Survey Telescope LSST](#) (8.4 m, 3200 MP, 3.5° wide field telescope, Cerro Pachón, Chile, first light 2019⁴)
- [NASA Today](#)
- [NASA/IPAC Extragalactic Database \(NED\)](#)
- <http://www.noao.edu/> National Optical Astronomy Observatory NRAO (by Assn. of Universities for Research in Astronomy)
- [N-Body Shop Galaxy simulations](#) (MPG movies and [home page](#))
- [Orbiting Binary Stars Radial Velocity Simulation](#) (needs JAVA, adjust M1, M2, a, e, i, w)
- [Planetary Nebula Spectra](#)
- [Pulsar Sounds](#)
- [Pulsation Mode in White Dwarf Simulation](#) (needs JAVA)
- [SIMBAD](#) ("Set of IDs, Measurements, and Bibliography For Astronomical Data" dataset, CDS=Centre de Données astronomiques de Strasbourg, France;) ♦♦
- [Sloan Digital Sky Survey SDSS](#) (star dataset w 2.5 m telescope, Apache Point Obs., NM) ♦♦
- [Smithsonian Astrophysical Observatory SAO](#)
- [Solar and Heliospheric Observatory SOHO](#) (including Current Solar Images) ♦♦
- [Space Telescope Science Institute](#) (including Hubble Space Telescope HST, 1990–)
- [Space Telescope Science Institute \(STScI\) Digitized Sky Survey \(DSS\)](#): includes [Digitized POSS](#) (Palomar Observatory Sky Survey), Hubble HST Phase 2
- [Spitzer Space Telescope](#) (2003–): one of NASA's 3 remaining Great Observatories in space, IR
- [Rotating Sky Explorer](#) (Flash animation)
- [United States Naval Observatory USNO Astronomical Applications](#) (Almanac)
- [UW Astronomy Department](#)

Space Weather (Earth–Solar Interactions)

- [Spaceweather.com](#) ♦
- [NOAA Space Weather Prediction Center](#) ♦
- [NOAA National Geophysical Data Center page on Solar-Terrestrial physics](#)

⁴ http://www.lsst.org/files/img/LSST_Timeline.jpg

Basic Physical Quantities for the Sun and the Earth

Symbols used are Sun = \odot and Earth = \oplus

Radius of Sun R_\odot

Equatorial Radius⁵ = 6.955×10^8 m = 695,500 km = 432,163 mi = 109 R_\oplus
Mean Diameter: 1.392×10^6 km

Radius of Earth (R_E or R_\oplus)⁶

Mean at Equator = 6378.1 km
Mean at Poles = 6356.8 km

Ellipticity of Sun f_\odot (Flattening or oblateness)⁷ = 9×10^{-6} (thus the Sun is a nearly perfect sphere)

Ellipticity of Earth = 0.00335 = ~0.3%

Distance from Sun to Earth, average (1 Astronomical Unit = 1 AU)⁸

1.496×10^8 km = 1.496×10^{11} m = 92.96×10^6 mi = 4.8481×10^{-6} parsec = $215 R_\odot$ = $108 D_\odot$
= 8.32 lm (light minutes) = 499 ls (light seconds)

Mass of Sun M_\odot ⁹ = 1.989×10^{30} kg = 333,000 M_\oplus

Mass of Earth M_\oplus = 5.9736×10^{24} kg

Density of Sun ρ_\odot [note that water density = 1 gm/cm³ = 1000 kg/m³]

Mean: 1.408 gm/cm³ = 1408 kg/m³ = 0.255 mean ρ_\oplus

Central or Core: up to 150 g/cm³ = 162×10^3 kg/m³ (the core extends to 0.2 to 0.25 R_\odot)

Photosphere¹⁰ = 2×10^{-7} g/cm³ = 2×10^{-4} kg/m³ [$\sim 1.6 \times 10^{-4}$ of Earth sea level atmos. density¹¹]

Particle density = $\sim 10^{23}$ particles m⁻³

Corona Mean¹² = 1×10^{-15} g/cm³ = 1×10^{-12} kg/m³

Lower Corona¹³ = 1×10^{-16} g/cm³

Particle density¹⁴ = $10^{15} - 10^{16}$ particles m⁻³

Density of Earth¹⁵ ρ_\oplus

Mean = 5.515 gm/cm³ = 5515 kg/m³

Inner Core = 12.8 – 13.1 gm/cm³ = 1280 – 1310 kg/m³

Mantle = 3.4 – 5.6 gm/cm³ = 3400 – 5600 kg/m³

Crust = 2.2 – 2.9 gm/cm³ = 2200 – 2900 kg/m³

Atmosphere (at sea level)¹⁶ = 1.2×10^{-3} gm/cm³ = 1.2 kg/m³

Temperature of Sun¹⁷

Surface effective black body temperature $T_{\text{eff},\odot}$ = 5777 K = 9,939 °F (IMA2, by Stefan–Boltzmann law)
Central¹⁸ or Core (modeled) = 15,700,000 K

⁵ Solar Radius:

• <http://solarscience.msfc.nasa.gov/> and
• <http://en.wikipedia.org/wiki/Sun>

⁶ http://en.wikipedia.org/wiki/World_Geodetic_System and <http://en.wikipedia.org/wiki/Earth>

⁷ Ellipticity:

<http://en.wikipedia.org/wiki/Flattening>:

Ellipticity is defined for an ellipsoid with equatorial radius a and polar radius b as $(a-b)/a$

⁸ http://en.wikipedia.org/wiki/Astronomical_unit

⁹ <http://nssdc.gsfc.nasa.gov/planetary/factsheet/sunfact.html>

¹⁰ <http://en.wikipedia.org/wiki/Photosphere>

¹¹ <http://solarscience.msfc.nasa.gov/>

¹² <http://en.wikipedia.org/wiki/Sun>

¹³ <http://solar-center.stanford.edu/vitalstats.html>

¹⁴ <http://en.wikipedia.org/wiki/Sun>

¹⁵ <http://en.wikipedia.org/wiki/Earth>

¹⁶ http://en.wikipedia.org/wiki/Earth's_atmosphere

¹⁷ <http://solarscience.msfc.nasa.gov/>

Luminosity of Sun L_{\odot} , Bolometric^{19,20} = $3.846 \times 10^{26} \text{ W} = 3.846 \times 10^{33} \text{ erg/s}$

Luminosity of Sun L_{\odot}^{Blue} (in the Blue-Band B)²¹ $\cong 1.9 \times 10^{26} \text{ W}$ (note: rarely used in comparisons and ratios)

Mass / Luminosity Ratio of Sun ($M_{\odot} / L_{\odot}^{\text{Bol}}$) = 5172 kg/W (calculated with above values).

Intensity of the Sun, Mean I_{\odot} ²² = $2.009 \times 10^7 \text{ W} \cdot \text{m}^{-2} \cdot \text{sr}^{-1}$

Stellar Classification of Sun, Spectral and Luminosity (Morgan–Keenan)^{23,24}: G2V

Magnitude of Sun:²⁵

Absolute Visual (M_V at 10 pc): +4.83

Apparent Visual (V = m_V): -26.74

Elemental Composition of Sun

Photosphere (by number % or ppm of nuclei)²⁶ = H 91.0%, He 8.9%; O 774 ppm, C 330 ppm,
Ne 112 ppm, N 102 ppm, Fe 43 ppm, Mg 35 ppm, Si 32 ppm, S 15 ppm

Photosphere (by relative mass)²⁷ = 70% H, 28% He, 2% “metals” (C, N, O, etc.)

Central (by relative mass)²⁸ = 35% H, 63% He, 2% “metals” (C, N, O, etc.)

Age of Sun²⁹ = $4.57 \times 10^9 \text{ yr}$

Age of Earth³⁰ = $4.54 \times 10^9 \text{ yr}$

Rate of Mass Conversion to Energy of Sun³¹ = $4.3 \times 10^9 \text{ kg s}^{-1}$

(i.e., the equivalent mass m=E/c² converted to energy to yield the Sun's bolometric luminosity)

Sun Mean Energy Production per Unit Mass³² = $0.194 \times 10^{-3} \text{ J/kg} = 8.1 \times 10^{-7} \text{ kcal (“human calories”)/kg}$

Rotation Periods of the Surface of the Sun:³³

Sidereal equatorial: 24.47 d (rotation period in days to same location in star frame)

Synodic equatorial: 26.24 d (rotation period in days to same apparent location as viewed from Earth)

Sidereal Carrington (~26 degrees latitude): 25.38 d

Synodic Carrington (~26 degrees latitude): 27.2753 d

¹⁸ ibid.

¹⁹ ibid.

²⁰ Luminosity:

http://en.wikipedia.org/wiki/Solar_luminosity:

Luminosity here is total bolometric (wide spectrum) photon radiant energy output, and does not include neutrino radiant energy, which adds $0.1 \times 10^{26} \text{ W}$

²¹ www.astro.sunysb.edu/aevans/PHY523/.../useful-definitions-pp.pdf

²² <http://en.wikipedia.org/wiki/Sun>

²³ Star Spectral classification:

http://en.wikipedia.org/wiki/Spectral_classification:

For the Sun classified as G2V in the Morgan–Keenan classification,

– Letter G indicates a yellowish star with surface T 5,200–6,000 K

– Number 2 indicates two tenths of the range between star spectral class G0 and start of adjacent class K0

– Roman Number V indicates the width of certain absorption lines, which correlates with star size—V indicates a Main Sequence dwarf star.

²⁴ http://en.wikipedia.org/wiki/Hertzsprung%E2%80%93Russell_diagram

²⁵ <http://en.wikipedia.org/wiki/Sun>

²⁶ <http://nssdc.gsfc.nasa.gov/planetary/factsheet/sunfact.html>

²⁷ <http://solarscience.msfc.nasa.gov/>

²⁸ ibid.

²⁹ <http://en.wikipedia.org/wiki/Sun> and <http://solarscience.msfc.nasa.gov/>

³⁰ http://en.wikipedia.org/wiki/Age_of_the_Earth

³¹ <http://nssdc.gsfc.nasa.gov/planetary/factsheet/sunfact.html>

³² ibid.

³³ http://en.wikipedia.org/wiki/Solar_rotation

Solar Rotation period (an empiric formula by latitude at the surface for data collected 1967 – 1987):³⁴

$$\omega = 14.713 - 2.396\sin^2(\varphi) - 1.787\sin^4(\varphi) \text{ , where}$$

ω = deg/d sidereal

φ = solar latitude (deg)

This gives sidereal days: $0^\circ \rightarrow 24.47$ d; $26^\circ \rightarrow 25.38$ d; $30^\circ \rightarrow 25.71$ d; $60^\circ \rightarrow 30.22$ d; $90^\circ \rightarrow 34.19$ d

Total Solar Irradiance (TSI, aka the “**Solar Constant**”) is the irradiance (W m^{-2}) at all wavelengths of photons (effectively from about 10,000 nm to about 10 nm) at exactly 1 A.U.³⁵ on a surface perpendicular to the incoming rays:

Mean is per NASA SORCE³⁶ about 1361 W m^{-2} (was previously said to be 1366 W m^{-2})

Instantaneous solar irradiance³⁷ varies due to changing Earth–Sun distance and solar fluctuations from $1,412$ to $1,321 \text{ W m}^{-2}$

Obliquity or Tilt of Sun's rotational axis³⁸ with respect to the ecliptic plane: ~ 7.25 degrees

Tilt of Sun's magnetic dipole field axis³⁹ with respect to rotation axis: up to $\pm (10^\circ \text{ to } 20^\circ)$

Sun rotation direction: Counterclockwise (when viewed from the north)

(this is the same direction that the planets including Earth rotate and orbit around the Sun).

Sun Escape Velocity: Escape velocity⁴⁰ in general is given by $v_e = (2GM/r)^{1/2}$, where M is the mass of the massive body being escaped from starting at distance from the body's center r, and v_e is the minimum speed required (ignoring drag) when propulsion ceases at that distance. For the Sun's surface, the idealized escape velocity is calculated at 617.5 km/s , whereas for the Earth's surface, it is calculated at 11.2 km/s . However, there appears to be a complex relationship between this simplistic solar escape velocity estimate and the actual velocities attained in solar wind. The latter is found to be below v_e at least using estimates from actual measurements near the Sun (at $4R_\odot - 7R_\odot$) using data from the Ulysses probe,⁴¹ and certainly solar wind speed is often $< v_e$ by the time the solar wind has decelerated in travelling to Earth orbit.

³⁴ Solar differential rotation:

Snodgrass, H. B. & Ulrich, R. K., “Rotation of Doppler features in the solar photosphere”, *Astrophysical Journal*, Part 1 vol. 351, 1990, p. 309–316

³⁵ ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SOLAR_IRRADIANCE/COMPOSITE.v2.PDF

³⁶ <http://en.wikipedia.org/wiki/Sunlight>

³⁷ http://en.wikipedia.org/wiki/Solar_irradiation

³⁸ <http://solarscience.msfc.nasa.gov/sunturn.shtml>

³⁹ http://www.nso.edu/press/newsletter/Tilted_Solar_Magnetic_Dipole.pdf

⁴⁰ http://en.wikipedia.org/wiki/Escape_velocity

⁴¹ Efimova AI et al. “Solar wind velocity measurements near the sun using Ulysses radio amplitude correlations at two frequencies”. *Advances in Space Research*. Volume 35, Issue 12, 2005, Pages 2189–2194

Basic Astronomical Terms, Concepts, and Tools (Chapter 1)

This is only a limited listing of some of the topics and definitions in this chapter.

Distance Measures

These are given with various degrees of rounding.

(Image depicts typical distance measurement using geometric parallax.)⁴²

Astronomical Unit AU

The earth–sun distance⁴³ = 1.4960×10^{11} m $\approx 1.50 \times 10^8$ km ≈ 93 million miles ≈ 8 lightmin

Lightyear or light-year or light year ly

The distance light travels in 1 Julian yr = 9.461×10^{12} km = 9.461×10^{15} m.
computed by = 365.25×86400 s $\times 299,792,458$ m/s = 9.46073×10^{15} m

Parsec (distance at which 1 AU subtends 1 second of arc) = 206,265 AU = 3.26 ly = 3.1×10^{13} km

Hubble Law, H_0 and h

$v = H_0 d$ where d = proper distance to a galaxy in Mpc, v = recession proper velocity in km s⁻¹

$$H_0 = 100h \text{ km s}^{-1} \text{ Mpc}^{-1}$$

where the WMAP 2010 value of the dimensionless $h = h_{\text{WMAP}} = 0.710^{+0.04}_{-0.03}$

Speed of Light Speed and Light Travel time/distances

c (speed of light) = 299,792,458 m/s (exactly, by definition) $\approx 3 \times 10^8$ m/s $\approx 3 \times 10^5$ km/s $\approx 186,282$ mi/s

1 lightsec = 3×10^5 km

1 lightmin = 18 million km (3×10^5 km/s \times 60s)

1 lightyear (ly, see above)

Representative distances

Earth Radius, equatorial mean = 6,377 km (diameter = 12,754 km)

Earth to moon distance, mean⁴⁴ $\approx 384,400$ km $\approx 239,000$ mi ≈ 1.3 lightsec

Sun to Pluto distance = 30 to 49 AU, 5 lighthours

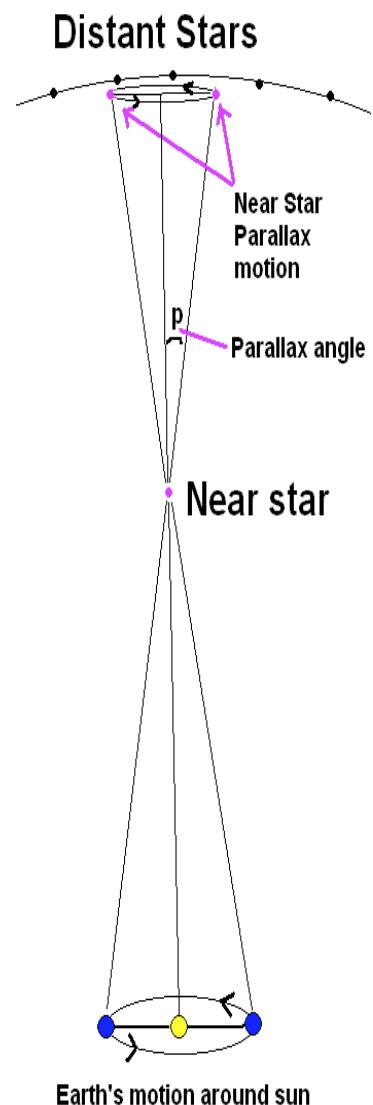
Sun to Oorts Cloud \approx 5,000 to 50,000 AU

Sun to Proxima Centauri (nearest star) = 4.24 LY

Sun to Large Magellanic Cloud (a nearby galaxy⁴⁵) ≈ 157 kly (48.5 kpc)

Milky Way disk diameter (visible) ≈ 50 kpc $\approx 163,000$ LY

Milky Way–Andromeda Galaxy ≈ 778 kpc



Distance Measured
by Parallax

⁴² <http://en.wikipedia.org/wiki/Parallax>

⁴³ http://en.wikipedia.org/wiki/Astronomical_unit

⁴⁴ http://en.wikipedia.org/wiki/Lunar_distance_%28astronomy%29

⁴⁵ http://en.wikipedia.org/wiki/Large_Magellanic_Cloud

Time Measures

See further below for sidereal time and epochs

Astronomical Year Numbering

"Astronomical year numbering is based on AD/CE year numbering, but follows normal decimal integer numbering more strictly. Thus, it has a year 0, the years before that are designated with negative numbers and the years after that are designated with positive numbers. Astronomers use the Julian calendar for years before 1582, including this year 0, and the Gregorian calendar for years after 1582... The year 1 BC/BCE is numbered 0, the year 2 BC is numbered -1, and in general the year n BC/BCE is numbered "-(n - 1)". The numbers of AD/CE years are not changed and are written with either no sign or a positive sign. The system is so named due to its use in astronomy... Although the absolute numerical values of astronomical and historical years only differ by one before year 1, this difference is critical when calculating astronomical events like eclipses or planetary conjunctions..."⁴⁶

Gregorian Calendar

"The Gregorian calendar, also called the Western calendar and the Christian calendar, is the internationally accepted civil calendar. It was introduced by Pope Gregory XIII (1502 - 1585), after whom the calendar was named, by a decree signed on 24 February 1582... The motivation for the Gregorian reform was that the Julian calendar assumes that the time between vernal equinoxes is 365.25 days, when in fact it is presently almost exactly 11 minutes shorter. The error between these values [resulted in an accumulated error of about 10 days by the time of the reform]..."⁴⁷ The rule for determining Leap Years now requires that a Leap Year (1)be evenly divisible by 4; (2) not evenly divisible by 100 unless also evenly divisible by 400. These rules have the effect of omitting 3 Leap Years every 400 years out of the 100 that are divisible by 4. The average [Gregorian] year length is $365 + (97/400) = 365.2425$ days per year, a close approximation to the tropical year of 365.2420 days(see below). The *proleptic Gregorian calendar* is produced by extending the Gregorian calendar backward to dates preceding its official introduction in 1582.⁴⁸

Heliocentric Julian Day HJD

is the Julian Date (JD) corrected for differences in the Earth's position with respect to the Sun. Even more precise corrections are made with the *Barycentric Julian Date*. "Due to the finite speed of light, the time an astronomical event is observed depends on the changing position of the observer in the Solar System. Before multiple observations can be combined, they must be reduced to a common, fixed, reference location. This correction also depends on the direction to the object or event being timed."⁴⁹ The BJD, which may differ from the HJD by up to 4 seconds, has replaced the HJD for precise time comparisons.

Julian Calendar

"The Julian calendar is a reform of the Roman calendar introduced by Julius Caesar in 46 BC (708 AUC). It took effect the following year, 45 BC (709 AUC), and continued to be used as the civil calendar in some countries into the 20th century. The calendar has a regular year of 365 days divided into 12 months.... A leap day is added to February every four years. The Julian year is, therefore, on average 365.25 days long.... The calendar year was intended to approximate the tropical (solar) year. Although Greek astronomers had known, at least since Hipparchus, that the tropical year was a few minutes shorter than 365.25 days, the calendar did not compensate for this difference. As a result, the calendar year gained about three days every four centuries compared to observed equinox times and the seasons. This discrepancy was corrected by the Gregorian reform, introduced in 1582."⁵⁰ The *proleptic Julian calendar* is produced by extending the Julian calendar backwards to dates preceding AD 4 when the quadrennial leap year stabilized.⁵¹

Julian Date JD

This is the number of days in the Julian Calendar since January 1, 4713 BCE at noon at Greenwich, i.e., by UT.

⁴⁶ http://en.wikipedia.org/wiki/Astronomical_year_numbering

⁴⁷ http://en.wikipedia.org/wiki/Gregorian_calendar

⁴⁸ http://en.wikipedia.org/wiki/Proleptic_Gregorian_calendar

⁴⁹ http://en.wikipedia.org/wiki/Barycentric_Julian_Date

⁵⁰ http://en.wikipedia.org/wiki/Julian_calendar

⁵¹ http://en.wikipedia.org/wiki/Proleptic_Julian_calendar

JD with fractional value of .0 is at noon UT
JD with fractional value of .5 is at midnight UT.
(Jan 1, 2012 at 0h UT = 2,455,927.5 JD and Jan 1, 2012 at noon UT is 2455928.0 JD)
See [here](#)⁵² for a JD calculator.

Modified Julian Day MJD

JD minus 2,400,000.5 day (used by spacecraft)

Note that with MJD, a fractional value of .0 is at midnight UT. One second after midnight, the MJD integer remains the same.

Second (SI)

The standard SI second is “the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium 133 atom”⁵³ There are $60 \times 60 \times 24 = 86,400$ SI seconds in the Julian day and $60 \times 60 \times 24 \times 365.25 = 31,557,600$ seconds in the Julian year. No other type of second is in common scientific usage, despite the confusion in days and years.



Terrestrial Time TT

“a modern astronomical time standard defined by the International Astronomical Union, primarily for time-measurements of astronomical observations made from the surface of the Earth... For example, the Astronomical Almanac uses TT for its tables of positions (ephemerides) of the Sun, Moon and planets as seen from the Earth. In this role, TT continues Terrestrial Dynamical Time (TDT),... which in turn succeeded ephemeris time (ET). The unit of TT is the SI second, the definition of which is currently based on the cesium atomic clock, but TT is not itself defined by atomic clocks. It is a theoretical ideal, which real clocks can only approximate. TT is distinct from the time scale often used as a basis for civil purposes, Coordinated Universal Time (UTC).”⁵⁴ TT takes into account effects of special and general relativity as Earth moves about the Sun and rotates. (*IMA2* p. 15)

Universal Time UT or GMT, UT1, UTC

local time at Greenwich (UT=PST+8 hr = PDT+7 hr). **UT1** is defined by Earth's actual rotation and thus based on irregular gravity effects. **UTC** is an acronym which stands for Universal Time Coordinated, a linguistic compromise between English Coordinated Universal Time and French temps universel coordonné. UTC is measured by atomic clocks, and adjusted by leap seconds as needed to keep UTC noon in synch with astronomical UT1 noon, etc.

Year

- The astronomical *Julian year* is exactly 365.25 days, each day of which has 86,400 SI seconds.
- The *tropical or solar year* is “the period of time for the ecliptic longitude of the Sun to increase by 360 degrees...The tropical year is often defined as the time between southern solstices, or between northward equinoxes [different values result]. Because of the Earth's axial precession, this year is about 20 minutes shorter than the sidereal year... The *mean tropical year*, as of January 1, 2000, was 365.241987 days [each of which has 86,400 SI seconds], or 365 days 5 hours, 48 minutes, 45.1897 seconds... According to Blackburn and Holford-Strevens (who used Newcomb's value for the tropical year) if the tropical year remained at its 1900 value of 365.24219878125 days the Gregorian calendar would be 3 days, 17 min, 33 s behind the Sun after 10,000 years... These effects will cause the calendar to be nearly a day behind in 3200.”⁵⁵

Long time intervals are measured in Ma = 10^6 Julian years, and Ga = Gyr = 10^9 Julian years, each year of which is 31557600 SI seconds in length. The Gy is the SI abbreviation for Gray, so this abbreviation is ambiguous for Gigayear. There is no SI abbreviation for year per se. Many persons use yr, such as Myr, Gyr. These abbreviations may be used for intervals. To express time ago I prefer Mya, Gya—e.g., the Big Bang occurred 13.75 Gya (13.75 Ga ago).

⁵² http://www.imcce.fr/en/grandpublic/temps/jour_julien.php

⁵³ <http://en.wikipedia.org/wiki/Second>

⁵⁴ http://en.wikipedia.org/wiki/Terrestrial_time

⁵⁵ partly from http://en.wikipedia.org/wiki/Tropical_year

Sky Position, Coordinate System, Motion and Rotation, Ecliptic, and Sidereal Terms

Altitude h

also called elevation, the angle above the local horizon measured along a great circle passing through the object and the zenith, 0 – 90°. The local altitude of the NCP is the same as local latitude in N hemisphere.

Azimuth A

angle in degrees measured E along the horizon starting at the North point [intersection of the meridian with the horizon that is closest to the NCP] to the intersection of the great circle used to measure altitude of the object (0 to 360°). Undefined at the Poles.

Celestial Equator CE

projection of the earth's equator onto the celestial sphere

Celestial Sphere

including *North Celestial Pole* NCP, *South Celestial Pole* SCP, & *Celestial Equator CE*

Declination δ

the angle N (+) or S (-) from the celestial equator for an object in sky

Ecliptic

The apparent path of Sun on the celestial sphere. More precisely, it is the path (the *Heliocentric ecliptic*) as seen from the Earth throughout the course of a year of the Sun's center. (Or at least it is the path if the celestial sphere and the Sun could be seen at the same time.) For potentially greater accuracy but more difficult to compute, the apparent path of the solar system's barycenter (center of gravity, a point which usually falls within the Sun, see below)⁵⁶ may be used to define the ecliptic.⁵⁷

Epoch

Positions of celestial objects must be specified in terms of a particular time or epoch, in order to account for precession, nutation, proper motion, etc. Currently used is epoch *J2000.0* (position at noon UT on January 1, 2000). An earlier system used Besselian years such as *B1950.0* as the time of reference.⁵⁸

Equatorial Coordinate System

The equatorial coordinate system is widely-used to map celestial objects. It projects the Earth's geographic poles and equator onto the celestial sphere. The projection of the Earth's equator onto the celestial sphere is called the *celestial equator*. Similarly, the projections of the Earth's north and south geographic poles become the north and south celestial poles, respectively. Coordinates are given as *Declination δ* and *Right Ascension RA*.

Equinox

"An *equinox* occurs twice a year, when the tilt of the Earth's axis is inclined neither away from nor towards the Sun, the center of the Sun being in the same plane as the Earth's equator..."⁵⁹ They are termed Vernal (March) equinox and autumnal (September) equinoxes... The equinoxes are currently in the constellations of Pisces and Virgo.

Heliocentric Model of Planetary Motion

Suggested by Nicolaus Copernicus (1473 – 1543) in his *De revolutionibus orbium coelestium*, published only in 1543, near the end of his life, out of fear. Elaborated and expanded by Johannes Kepler (1571 - 1630) in his *Mysterium Cosmographicum*, published in 1596, and by Galileo Galilei (1564 – 1642) in his 1610 *Sidereus Nuncius* (Starry Messenger) and his 1632 *Dialogo sopra i due massimi sistemi del mondo* (*Dialogue Concerning the Two Chief World Systems*).

Hour Angle HA

Angle W along CE from meridian to hour circle

Hour Circle

The great circle through an object and the celestial poles, therefore perpendicular to the CE.

⁵⁶ <http://en.wikipedia.org/wiki/Ecliptic>

⁵⁷ <http://www.astro.sunysb.edu/fwalter/PHY515/glossary.html>

⁵⁸ http://en.wikipedia.org/wiki/Epoch_%28astronomy%29#Besselian_years

⁵⁹ <http://en.wikipedia.org/wiki/Equinox>

Horizon

Where the sky meets the ground (in alt-azimuth coordinate system)

Horizon, Horizontal, or Alt/Az Coordinate System

Uses local horizon to measure:

altitude (0 – 90° above the horizon, also called elevation)

azimuth (0 – 360° measured from North Point on the horizon toward the E, i.e., to the right)

Meridian

Great circle through local zenith and NCP and SCP. The Prime Meridian is the meridian the runs through Greenwich, England (0 degrees longitude).

Planetary Positions

Positional terminology expressed relative to the Earth and Sun differ depending on their orbits relative to the Earth:

Inferior planet: these orbit inside the orbit of the Earth (Mercury, Venus). Relative to Earth, these assume greatest eastern elongation, greatest western elongation, superior conjunction (in same direction as Sun but beyond Sun), inferior conjunction (in same direction as Sun but closer than Sun).

Superior planet: these orbit outside that of the Earth (Mars, Jupiter, Saturn, Uranus, Neptune). Relative to Earth, these assume eastern quadrature, western quadrature, conjunction (in same direction as Sun but beyond Sun), and opposition (on an extended Earth Sun line in direction opposite to the Sun).

Precession

Arises from tidal forces or torques caused by gravitational forces by Sun and moon on the oblate and tilted Earth. Earth's axial tilt varies between 22.1° and 24.5°. Earth precesses over a period of 25770 y, with current axial tilt angle 23.5° or 23.4°. Precession necessitates correction of right ascension and declination of astronomical bodies and equinoxes etc. relative to the standard reference time when their positions are tabulated, for instance to epoch J2000.0 (Noon UT in Greenwich England on Jan 1, 2000). (IMA2 p. 13)

Proper Motion μ and Velocities

Proper Motion is the apparent angular motion transverse to the line of sight between the observer and a celestial body at distance r . It is expressed as an angular velocity $\mu = d\theta/dt = v_\theta/r$, expressed typically for stars in arcsec/yr. The linear velocity of the object \mathbf{v} is decomposed into the *transverse linear velocity* v_θ and the *radial linear velocity* v_r . An angular distance traveled may be expressed in terms of changes in RA and declination by (IMA2 p. 19):

$$(\Delta\theta)^2 = (\Delta\alpha \cos \delta)^2 + (\Delta\delta)^2$$

Retrograde Motion

Apparent retrograde motion is the motion of a planetary body in a direction opposite to that of other bodies within its system as observed from a particular vantage point. Direct motion or prograde motion is motion in the same direction as other bodies.⁶⁰ Retrograde motion was difficult to explain in the Ptolemaic geocentric universe (named for Claudius Ptolemaeus, c. 90 CE – 168 CE, an astronomer who brought the geocentric model to its near-highest form).

Right Ascension RA or α

The angle E along the celestial equator from Υ to the hour circle of a star. It is the angle measured eastward along the CE from the vernal equinox to its intersection with the object's hour circle (IMA2 p. 12) It is traditionally measured in hours, minutes, and seconds.

Rotation Direction

When referring to the Earth and the Celestial sphere (see *here*⁶¹ for animation):

From a Fixed CE Perspective: If the Earth is rotating as viewed looking inward from outside a stationary Celestial Sphere, the rotation of points on the Earth is to the right (CCW looking down from NCP, CW down from the SCP) and this rightward direction is toward the Earth's east or eastward.

From a Fixed Earth Perspective: If one is on a fixed earth looking out at a rotating Celestial Sphere in the N hemisphere facing N, the visible CS appears to rotate CCW about the NCP (CW about the SCP in the S hemisphere), with non-circumpolar objects rising on the right in the east and rotating up and toward the left then descending in the west. Looking toward the S in the N hemisphere, stars rise on the left (east) and rotate up and set on the right (west).

⁶⁰ http://en.wikipedia.org/wiki/Apparent_retrograde_motion

⁶¹ <http://astro.unl.edu/classaction/animations/coordsmotion/celhorcomp.html>

Sidereal Period

The amount of time that it takes an object to make one full orbit, relative to the stars. This is considered to be an object's true orbital period. Compare synodic period.

Sidereal Time ST

RA + HA. Local sidereal time is the amount of time that has elapsed since the vernal equinox last traversed the (local) meridian. A mean sidereal day is shorter than a mean solar day, namely

$$\sim 23\text{h } 56\text{m } 04.\text{s} = 86164 \text{ s}$$

vs.

$$24^{\circ}60'60'' = 86400 \text{ s}$$

in the mean solar (tropical) day. Thus the mean sidereal day is shorter by 236 s. This is calculated by $86,400\text{s} \times 365.242190402 / 366.242190402 = 86164 \text{ s}$, because there are 1 more sidereal days than tropical days in each tropical year.⁶² "The right ascension of a star is equal to the sidereal time when that star crosses the meridian.... By observation, you could find the sidereal time by locating a prominent star that is passing directly overhead (crossing the meridian) and then looking up its right ascension on your star finder or in a star atlas.... For example, the star Capella in the constellation Auriga has the coordinates RA 5h 13m DEC 46deg. When Capella is directly overhead, the sidereal time is 5:13."⁶³ See [here](#)⁶⁴ for accurate determination of sidereal time.

Solstice

"A solstice is an astronomical event that happens twice each year when the Sun's apparent position in the sky reaches its northernmost or southernmost extremes. The name is derived from the Latin *sol* (sun) and *sistere* (to stand still), because at the solstices, the Sun stands still in declination; that is, the apparent movement of the Sun's path north or south comes to a stop before reversing direction."⁶⁵ They are termed Summer (or less ambiguously Northern or June) Solstice and Winter (or less ambiguously Southern or December) solstices.

Their precise times in 2012 in UT are given [here](#):⁶⁶

Perihelion	Jan 4 at 2:08	Equinoxes	Mar 20 at 5:20	Sept 22 at 15:01
Aphelion	July 4 at 17:03	Solstices	June 20 at 23:21	Dec 21 at 11.23

Synodic period

For planets relative to the Earth, the time interval between successive oppositions or conjunctions, etc. "The synodic period is the temporal interval that it takes for an object to reappear at the same point in relation to two other objects (linear nodes), e.g., when the Moon relative to the Sun as observed from Earth returns to the same illumination phase. The synodic period is the time that elapses between two successive conjunctions with the Sun-Earth line in the same linear order. The synodic period differs from the sidereal period due to the Earth's orbiting around the Sun."⁶⁷

Vernal Equinox γ

The point on the CE where the path of the Sun (ecliptic) intersects or crosses from S to N. (The symbol γ represents Aries, Latin for *ram*, not the Greek letter gamma). Also called the *First Point of Aries* [though now actually falling in constellation Pisces], aka *the ascending node*. Also represents the time of this crossing, typically falling between 3/19 to 3/21. The autumnal equinox falls between Sept. 22 to 24.

Zenith

The point directly over the observer's head. In alt-az coordinate system, this is alt = 90.

⁶² http://en.wikipedia.org/wiki/Sidereal_time

⁶³ <http://astronomy.uconn.edu/defs/sidereal.html>

⁶⁴ <http://aa.usno.navy.mil/faq/docs/GAST.php>

⁶⁵ <http://en.wikipedia.org/wiki/Solstice>

⁶⁶ <http://aom.giss.nasa.gov/cgi-bin/srevents.cgi>

⁶⁷ http://en.wikipedia.org/wiki/Synodic_period

Celestial Mechanics (Chapter 2)

This chapter has not yet been read in my astronomy courses and only parts of it are summarized here.

Kinematics Defined

“Kinematics is the branch of classical mechanics that describes the motion of points, bodies (objects) and systems of bodies (groups of objects) without consideration of the forces that cause it... The study of kinematics is often referred to as the geometry of motion... To describe motion, kinematics studies the trajectories of points, lines and other geometric objects and their differential properties such as velocity and acceleration.”⁶⁸

Mathematical Properties of the Ellipse

An ellipse is the set of points satisfying

$$r + r' = 2a$$

where a = *semimajor axis* (1/2 the length $2a$ of the long axis), and r and r' are the distances of any point on the ellipse from the two foci F and F' , respectively.

The *eccentricity* e ($0 \leq e < 1$) is defined as

$$e \equiv (\text{distance between the foci } F \text{ and } F')/2a.$$

The eccentricity is 0 for a circle. The distance from a focal point to the center of the ellipse (midway between the two foci) is ae . The *semiminor axis* is b .

Kepler's Original Laws of Planetary Motion

Johannes Kepler (1571 – 1630) published his first two laws of Planetary Motion in 1609 and his third in 1619.

Kepler's First Law

A planet orbits the Sun in an ellipse, with the Sun at one focus. The closest approach of a planet to the Sun is on the major axis and is termed the **perihelion**, the furthest being the **aphelion**.

The distance on an ellipse from the principal focus is given in polar coordinates by

$$r = a(1 - e^2)/(1 + e \cos \theta) \text{ where } (0 \leq e < 1)$$

and where θ is the angle measured CCW from the perihelion position (which is taken to be 0).

Kepler's Second Law

A line connecting a planet to the Sun sweeps out equal areas in equal time intervals.

Kepler's Third Law

$$P^2 \propto a^3$$

where P is the orbital period of the planet, and a is the average distance of the planet from the Sun.

For solar planets, if P is measured in sidereal years, and a measured in AU,

$$P^2 = a^3$$

⁶⁸ <http://en.wikipedia.org/wiki/Kinematics>

Laws of Gravity and Motion

Newton's Law of Universal Gravitation

Isaac Newton (1642 - 1727) published his three laws of motion and the law of universal gravitation in *Philosophiae Naturalis Principia Mathematica* in 1687:

$$F = GMm/r^2$$

where F is the gravitational force between two masses, G the universal gravitational constant ($6.673 \times 10^{-11} \text{ N m}^2 \text{ kg}^{-2}$), M and m are the masses of the two objects, and r is the distance between them. For extended masses, F is determined by integrating over the mass distribution(s). For spherically symmetrical mass distributions, F is directed along the line of symmetry between the two objects and acts as if all the mass of the spherically symmetrical object were located at the center of the mass. *IMA2* proves that, "For points inside a spherically symmetrical distribution of matter, the portion of the mass that is located at radii $r < r_0$ causes the same force at r_0 as if all of the mass enclosed within a sphere of radius r_0 was concentrated at the center of the mass distribution. However, the portion of the mass that is located at radii $r > r_0$ exerts no net gravitational force at the distance r_0 from the center."⁶⁹

For such a spherically symmetrical body such as a planet, the *escape velocity* is given by

$$v_{\text{esc}} = (2GM/r)^{1/2}$$

where M is the mass of the planet to be escaped from and r is the distance from the center of that planet.

Kepler's Laws Updated (by Isaac Newton, etc.)

Define *reduced mass* μ as

$$\mu = m_1m_2/(m_1 + m_2)$$

The reduced mass is used in a center-of-mass reference frame in order to reduce the problem to a rotation of a reduced mass about the total mass located at the origin.

Define θ as the angle of the reduced mass as measured from the direction to perihelion.

The orbital angular momentum L is conserved in binary orbits.

Kepler First Law, revised: Both objects in a binary orbit move about the center of mass in (separate) ellipses, with the center of mass occupying one focus of the ellipse. In a center-of-mass reference frame, if L is the angular momentum (units of $\text{kg}\cdot\text{m}^2 \text{ s}^{-1}$), then the distance r of the reduced mass from the total mass is given by

$$r = (L^2/\mu^2) / GM(1 + e \cos \theta)$$

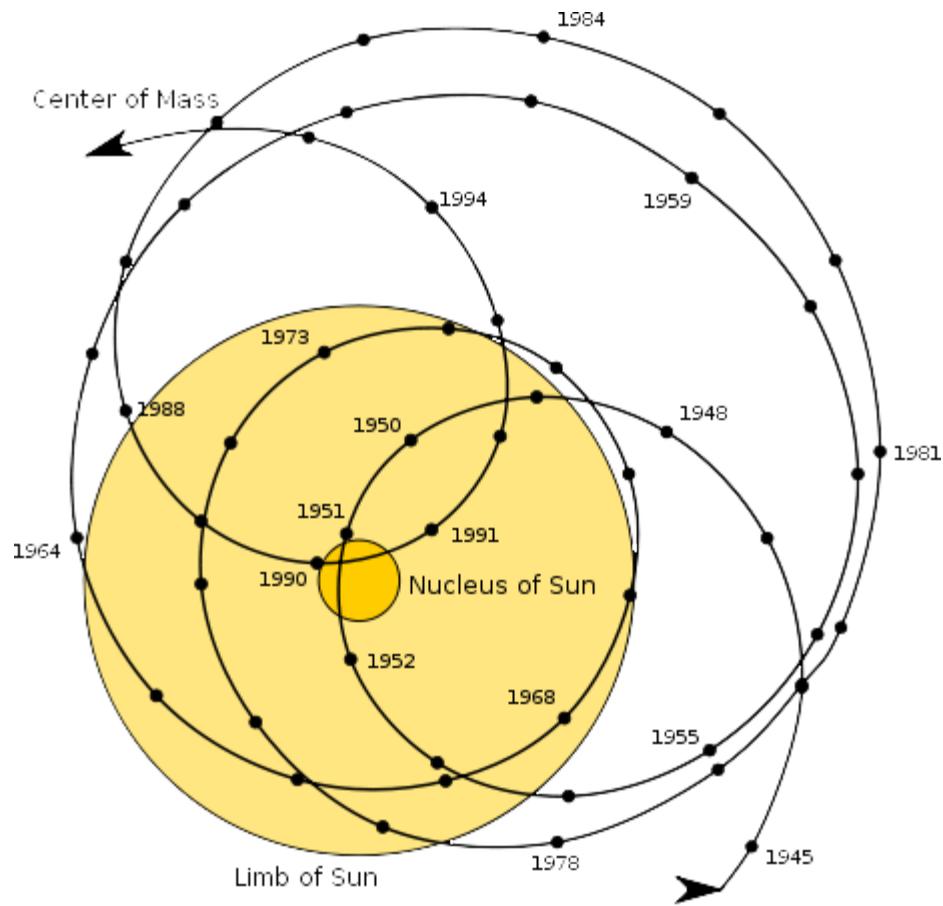
where M is the total mass of the system, e is the redefined eccentricity, and θ is as above.

This is the equation of a conic section, an ellipse if the total energy of the system is less than zero (i.e., a bound system).

In multiple body systems, the barycenter (center of gravity) moves in a more complex manner. For example, the motion of the barycenter of the Solar System relative to the center of the Sun is depicted in the following graph, which spans 50 years.⁷⁰ Note that the barycenter has not always fallen inside the Sun surface.

⁶⁹ http://en.wikipedia.org/wiki/Newton%27s_law_of_universal_gravitation

⁷⁰ <http://en.wikipedia.org/wiki/Sun#Atmosphere>



Path of Solar System Center of Mass (Barycenter) Relative to the Sun Center, 1945 - 1995

Kepler Second Law, revised:

$$dA/dt = (L/\mu)/2$$

where the right hand side is constant. Presumably (I have not confirmed this) the area is swept out by the reduced mass and is measured from the center of mass.

Kepler Third Law, revised:

$$P^2 = [4\pi^2 / G(m_1 + m_2)] a$$

Virial Theorem

The Virial theorem was named in 1870 by Rudolf Clausius (1822 – 1888), in German fr. Latin, *vir* = plural stem of *vis*: strength, force). For a system that is stably bound by gravitational forces (or by other inverse-square-law binding forces):

$$\langle E \rangle = \langle K \rangle + \langle U \rangle = 1/2 \langle U \rangle$$

where E is the total energy, $\langle E \rangle$ is the time-averaged total energy, $\langle K \rangle$ is the time-averaged kinetic energy KE, and $\langle U \rangle$ is the time-averaged potential energy PE. As with atoms, the total energy for a gravitationally bound system is considered to be negative, inasmuch as the components are considered to have zero energy when infinitely separated, and energy must be radiated away or emitted when the components accelerate inward to form a bound system. A stably bound system for which the virial theorem applies is said have undergone **virialization** and to have become "**virialized**". There are scalar and tensor forms of this useful theorem (and it has been extended to deal with electric and magnetic fields). It has been used to deduce the presence of dark matter.

Light Properties (Pre-QM and Early QM); Distance and Magnitudes (Chapter 3)

Astronomical Distance, Luminosity, Magnitude

Stellar Parallax

Astrometry is the branch of astronomy that is concerned with precise determination of 3-dimensional positions of celestial objects, including their distance.

Stellar parallax is a method of measuring distance to a star (or other celestial object) that is sufficiently nearby (see image in Chapter 1). The parallax angle p (seconds of arc) is formed by the angle between the sun, the star, and the earth (note: not the angle formed with the Sun and Earth at opposite orbital locations). For small angles, distance d is given by

$$d \text{ (in pc)} = 1/p \quad \text{where } p \text{ is in arcseconds}$$

$$d \text{ (in AU)} = 206,265/p \quad \text{where } 206,265 = \text{arcseconds in a radian}, p \text{ is in arcseconds}$$

By definition, if $p = 1$, $d = 1$ parsec. The nearest star, Proxima Centauri, at 1.3 pc or 4.2 ly has a parallax angle of only 0.77 arcseconds. The Hipparcos Space Astrometry Mission (HSAM, 1989–1993) was able to measure some parallaxes as small as 0.6 mas corresponding to 1600 ly or 500 pc. More accurate missions are anticipated—GAIA, launching 2013, is expected to attain “an accuracy down to 20 μ as”.⁷¹

Luminosity L

is the total energy (J) emitted as electromagnetic radiation or photons (“light”) per second (thus expressed in W or J s^{-1}). Solar luminosity L_\odot varies but is about $3.839 \times 10^{26} \text{ W}$. (Slightly more energy is emitted if neutrinos are also included. Particles in the solar wind are not included.) When the term is not otherwise qualified, luminosity is the same as total *bolometric luminosity*. (A *bolometer* is an instrument that measures radiant energy over a wide band by absorption and measurement of heating.)

Radiant (or radiative) flux F or S

Light from a star emitting uniformly in all directions varies as the inverse square of the distance r (*inverse square law*), and is given by $F = S = L/4\pi r^2$. (The inverse square law does not apply to a non-spherically-symmetrical or collimated light beam.) Flux is expressed as power per unit area (W m^{-2}) passing through a plane that is (typically) perpendicular to the direction of radiation travel.

Total Solar irradiance (flux) at 1 A.U. (at the outer Earth atmosphere) varies but averages about 1365 W m^{-2} (see later in this document).

Specific Flux density S_ν may also be expressed in Janskys ($10^{-26} \text{ watts m}^{-2} \text{ Hz}^{-1}$), named for Karl Guthe Jansky (1905 – 1950, one of the pioneers of radio astronomy). Jansky is a non-SI unit approved by the IAU=International Astronomical Union, and widely used in radio and infrared astronomy. Specific Flux density incorporates the Hz^{-1} unit, and therefore Specific Flux density must be integrated over the finite receiving band of wavelengths being detected to yield the total flux in this band.⁷²

⁷¹ <http://sci.esa.int/science-e/www/object/index.cfm?fobjectid=31197>

⁷² IRA3 p. 9 = Burke BF and Graham-Smith F, *An Introduction to Radio Astronomy*, 3rd Ed. Cambridge. 2010
Page 18 of 249

Magnitude and Distance Modulus

Apparent magnitude m (m_{bol})

This is the apparent brightness of an object, in which a plus 1 unit increment represents a brightness reduction by a factor of $100^{1/5} = 2.512$. The incremental star magnitudes scale derives from the ancient Greeks, who categorized magnitudes in 6 groups. Smaller (less positive or more negative values) represent brighter objects. Apparent magnitude here is *bolometric* (thus measured over all wavelengths of light) and indicated with m_{bol} , but the term is also used to measure apparent color magnitude (e.g., in blue light, which is abbreviated m_B). The m_{bol} for the Sun is -26.83 , and $m = +30$ for the faintest detectable objects, representing a ratio of brightness of about 10^{23} .

The zero point (zeropoint) of magnitude is approximately the magnitude of Alpha Lyrae or Vega⁷³. Specifically, in the Johnson system, the “zero” point is chosen such that $V = 0.03$ for Alpha Lyrae (Vega). But current definitions for zero point vary and are complicated. For instance, HST states for its Wide-Field Planetary Camera 2 or WFPC2, “The zeropoints in the WFPC2 synthetic system, as defined in Holtzman et al. (1995b), are determined so that the magnitude of Vega, when observed through the appropriate WFPC2 filter, would be identical to the magnitude Vega has in the closest equivalent filter in the Johnson-Cousins system. For the filters in the [WFPC2] photometric filter set, F336W, F439W, F555W, F675W, and F814W, these magnitudes are 0.02, 0.02, 0.03, 0.039, and 0.035, respectively.”⁷⁴

For any two magnitudes m_1 and m_2 (bolometric or color), the ratio of their fluxes is given by

$$F_2/F_1 = 100^{(m_1 - m_2)/5}$$

Absolute magnitude M

This is the calculated apparent bolometric (or color) magnitude of an object at a distance of 10 pc.

Distance Modulus m - M

The distance to a celestial object is given by

$$d = 100^{(m - M + 5)/5} \text{ pc}$$

where m and M are apparent and absolute magnitudes. The quantity $(m - M)$ therefore serves as a measure of distance, and is given by

$$m - M = 5 \log_{10}(d) - 5 = 5 \log_{10}(d/10 \text{ pc})$$

where d is in pc

Comparison with the Sun

The absolute magnitude of a star expressed in terms of luminosity of it and the Sun is

$$M = M_{\text{Sun}} - 2.5 \log_{10}(L/L_{\odot})$$

where $M_{\text{bol,Sun}} = 4.75$ and $L_{\odot} = 3.846 \times 10^{26} \text{ W}$. Here, M is magnitude, not mass.

⁷³ Michael S. Bessell. *Annu. Rev. Astron. Astrophys.* 2005. 43:293–336

⁷⁴ http://www.stsci.edu/instruments/wfpc2/Wfpc2_dhb/wfpc2_ch52.html

Surface Brightness

For extended objects (such as galaxies, star clusters, or nebulas) that are not point-sources of light, it is customary to express their **surface brightness** S (or I) quoted in units of magnitude per square parsec (i.e., mag arcsec $^{-2}$). The word “surface” refers to the fact that the light originates from a spread-out surface on part of the celestial sphere rather than from a point. For the same emitting object, surface brightness does not change with increasing distance (ignoring extinction, etc.).... Since the light is spread out, the average brightness from any point on this surface is generally much fainter than a star as a point source at comparable distance.⁷⁵

Because the magnitude is logarithmic, calculating surface brightness cannot be done by simple division of magnitude by area. Instead, for a source with magnitude m extending over an area of A in square arcseconds, the surface brightness S (quoted in units of magnitude per square arcsecond or mag arcsec $^{-2}$) is given by⁷⁶

$$S = m + 2.5 \log_{10} A$$

Surface brightness S is constant with luminosity distance. For nearby objects, the luminosity distance is equal to the physical distance of the object. For a nearby object emitting a given amount of light, radiative flux decreases [by the inverse] square of the distance to the object, but the physical area corresponding to a given solid angle (e. g., 1 arcsec) increases [also by the square of the distance]..., resulting in the same surface brightness. The relationship with standard solar magnitude and luminosity units is given by

$$S \text{ (in magnitude per arcsecond}^2\text{)} = M_{\odot} + 21.572 - 2.5 \log_{10} S \text{ (in } L_{\odot}/\text{pc}^2\text{)}$$

where again the left-hand S of the object is expressed in units of magnitudes per square arcsecond, M_{\odot} is the absolute magnitude of the Sun in the chosen color band, and the right-hand S of the object is expressed in units of luminosity of the Sun L_{\odot} [in the same color band?—this appears to be a point of confusion] per pc 2 .

“A great deal of confusion ensues from the fact that amateur astronomers habitually fail to specify whether they mean integrated brightness or surface brightness when they say that an object is bright or faint.

Consider, for instance, M33, the Triangulum Galaxy. At magnitude 5.7, it is fifth in integrated brightness of any galaxy in the sky, after our own Milky Way ... Nonetheless, M33 is referred to as a faint galaxy, because its light is spread out over a huge area—nearly a square degree—giving it one of the lowest surface brightnesses of any Messier object. On the other hand, the planetary nebula M76 has one of the highest surface brightnesses of any nebulous Messier object, but it is often called faint because of its low integrated brightness. (For mathematicians, the term **integrated brightness** refers to the integral of the surface brightness over the object’s area, which in the case of M76, is tiny.) ”⁷⁷

Light (Photon) Wave vs. Particle Properties

Astronomers tend to use the term “light” to refer to electromagnetic radiation of any wavelength. Light travels in vacuum at speed $c = 299,792,458$ m/s (exactly, by definition). Thomas Young (1773 – 1829) performed his double-slit diffraction experiment c. 1803, in which he showed that light constructively and destructively interferes, thereby establishing its wavelike behavior and deducing a *wavelength* λ . Given the finite speed of light c , the light *frequency* ν is given by

$$c = \lambda\nu$$

He determined the wavelengths of *visible light* (modern range = 390 to 750 nm, commonly stated as 400 to 700 nm, or 4000 to 7000 Å). The range 390 to 750 nm corresponds to frequencies of 400–790 THz (1 THz = 10^{12} Hz).

James Clerk Maxwell (1831 - 1879) used his Maxwell equations in 1865 to predict electromagnetic transverse waves propagating at

⁷⁵ <http://www.astro.lsa.umich.edu/undergrad/labs/brightness/index.html>

⁷⁶ http://en.wikipedia.org/wiki/Surface_brightness

⁷⁷ <http://mysite.verizon.net/vze55p46/id18.html>

$$v = 1 / (\epsilon_0 \mu_0)$$

where ϵ_0 = *permittivity of free space* (applying to electric fields) and μ_0 = *permeability of free space* (applying to magnetic fields). These were consistent with light waves.

The *Poynting Vector S* specified the classical (non-quantum-mechanical) energy (W m^{-2}) carried by a light wave:

$$\mathbf{S} = (1/\mu_0) \mathbf{E} \times \mathbf{B}$$

and the vector points in the direction of propagation. The time-averaged value (not root mean square) is

$$\langle S \rangle = (1/2\mu_0) E_0 B_0$$

although this applies to a specific frequency.

The EM wave carries momentum which can impart a force (i.e., it can exert *radiation pressure*), given by

$$F_{\text{rad}} = \langle S \rangle (A/c) \cos^2 \theta \quad \text{for absorption, or}$$

$$F_{\text{rad}} = 2\langle S \rangle (A/c) \cos^2 \theta \quad \text{for reflection}$$

where A is the area of the surface radiated and θ is the angle of incidence measured wrt perpendicular to the plane.

Typically, the radiation pressure arises from Compton scattering of photons by electrons, *IMA2* p. 119, and plays a significant role in the interiors of extremely luminous objects such as early main-sequence stars, red supergiants, accreting compact stars, and possibly on small dust particles in the ISM. The solar irradiance at the Earth exerts a pressure of about $4.6 \mu\text{Pa}$ (absorbed)⁷⁸ and its effects can be demonstrated with a highly sensitive Nichols radiometer (1903), but not with the Crookes radiometer (1873).

⁷⁸ http://en.wikipedia.org/wiki/Radiation_pressure

Spectrum of Electromagnetic Radiation (Photons of “Light”)⁷⁹

Spectrum of Electromagnetic Radiation (“Light”)

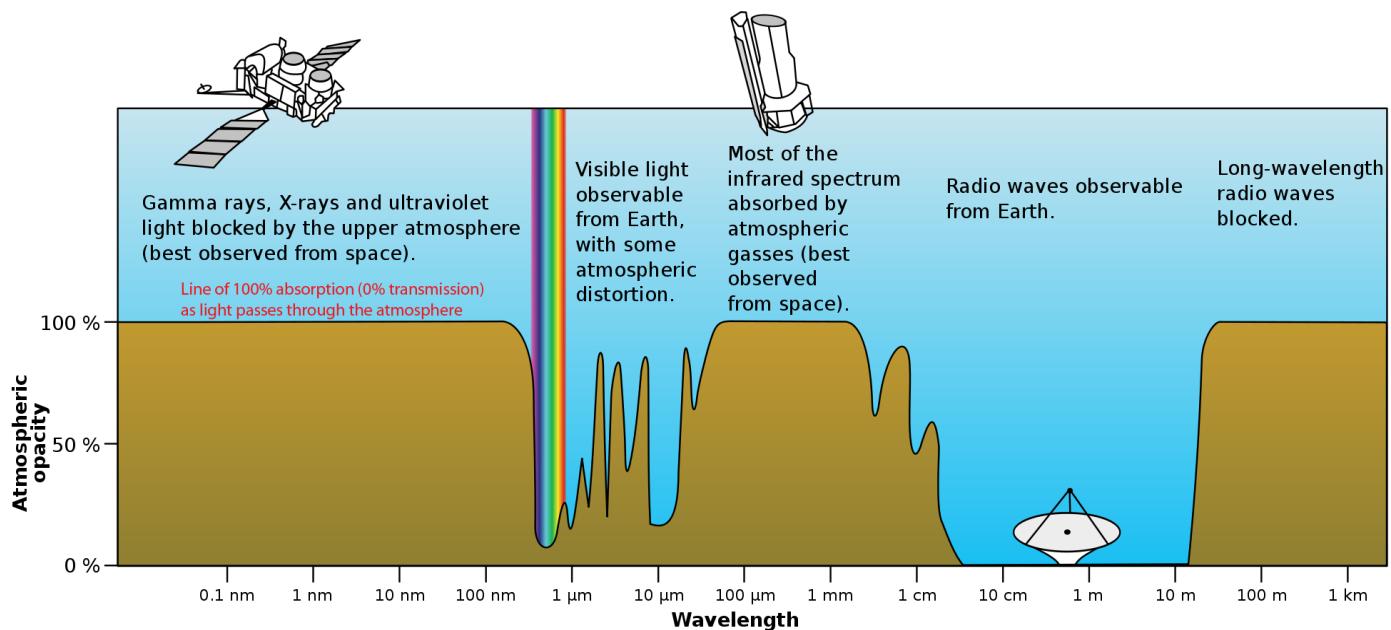
Region	Wavelength (nm)	Wavelength (m)	Wavelength (cm)	Frequency (Hz)	Energy (eV)
Radio (and longer λ)	$> 10^8$	> 0.1	> 10	$< 3 \times 10^9$	$< 1.24 \times 10^{-5}$
Microwave	$10^8 - 10^5$	$0.1 - 0.0001$	$10 - 0.01$	$3 \times 10^9 - 3 \times 10^{12}$	$1.24 \times 10^{-5} - 0.0124$
Infrared	$10^5 - 700$	$0.0001 - 7 \times 10^{-7}$	$0.01 - 7 \times 10^{-5}$	$3 \times 10^{12} - 4.3 \times 10^{14}$	$0.0124 - 1.77$
Visible Light	$700 - 400$	$7 \times 10^{-7} - 4 \times 10^{-7}$	$7 \times 10^{-5} - 4 \times 10^{-5}$	$4.3 \times 10^{14} - 7.5 \times 10^{14}$	$1.77 - 3.1$
Ultraviolet	$400 - 10$	$4 \times 10^{-7} - 10^{-8}$	$4 \times 10^{-5} - 10^{-6}$	$7.5 \times 10^{14} - 3 \times 10^{16}$	$3.1 - 124$
X-Rays (Soft & Hard)	$10 - 0.01$	$10^{-8} - 10^{-11}$	$10^{-6} - 10^{-9}$	$3 \times 10^{16} - 3 \times 10^{19}$	$124 - 1.24 \times 10^5$
Gamma Rays (of any origin)	< 0.01	$< 10^{-11}$	$< 10^{-9}$	$> 3 \times 10^{19}$	$> 1.24 \times 10^5$

Note: opinions differ as to the boundary between soft X-rays and extreme UV, somewhere between 1 & 10 nm. Many astronomers seem to use 10 nm. Distinction between radio and microwaves also varies.

⁷⁹ Electromagnetic Spectrum:

- <http://csep10.phys.utk.edu/astr162/lect/light/spectrum.html> (adapted & revised by MCM)
- <http://en.wikipedia.org/wiki/X-ray>

Earth's Atmospheric Absorption of Incoming EM Radiation ("Light")



Brown regions represent amount of attenuation of the wavelength that has occurred at ground level.
Greatest transparency is seen in visible light, some infrared bands, and at radio wavelengths.⁸⁰

Blackbody Radiation and the Planck Distribution Functions

A blackbody is an idealized perfect absorber of incident EM radiation, regardless of frequency or angle of incidence. "A black body in thermal equilibrium (that is, at a constant temperature) emits electromagnetic radiation called black-body radiation. The radiation is emitted according to Planck's law, meaning that it has a spectrum that is determined by the temperature alone..., not by the body's shape or composition... A black body in thermal equilibrium has two notable properties: (1) It is an ideal emitter: it emits as much or more energy at every frequency than any other body at the same temperature, and (2) It is a diffuse emitter: the energy is radiated isotropically, independent of direction.. An approximate realization of a black body is a hole in the wall of a large enclosure... Any light entering the hole is reflected indefinitely or absorbed inside and is unlikely to re-emerge, making the hole a nearly perfect absorber. The radiation confined in such an enclosure may or may not be in thermal equilibrium, depending upon the nature of the walls and the other contents of the enclosure. Real materials emit energy at a fraction—called the emissivity—of black-body energy levels. By definition, a black body in thermal equilibrium has an emissivity of $\epsilon = 1.0$. A source with lower emissivity independent of frequency often is referred to as a gray body. Construction of black bodies with emissivity as close to one as possible remains a topic of current interest... In astronomy, the radiation from stars and planets is sometimes characterized in terms of an effective temperature, the temperature of a black body that would emit the same total flux of electromagnetic energy."⁸¹

The *spectral or specific brightness* $B_\lambda(T)$ (called "the Planck function" and otherwise unnamed in IMA2, aka *spectral radiance* or *specific intensity* or *spectral intensity*) for values to be integrated over wavelength and solid angle differentials is expressed in units of $\text{W m}^{-2} \text{nm}^{-1} \text{sr}^{-1}$. Note that instead of using pure SI length units of m^{-3} , the nm^{-1} term corresponds to the $d\lambda$ wavelength differential to be integrated, whereas the m^{-2} term corresponds to the area of the emitter. $B_\lambda(T)$ is reserved for brightness at the emitter surface, whereas Specific Intensity I_λ , also expressed as $\text{W m}^{-2} \text{nm}^{-1} \text{sr}^{-1}$, is used for the local value of flux along a ray path, since it may change with emission and absorption processes (IRA3 p. 9).

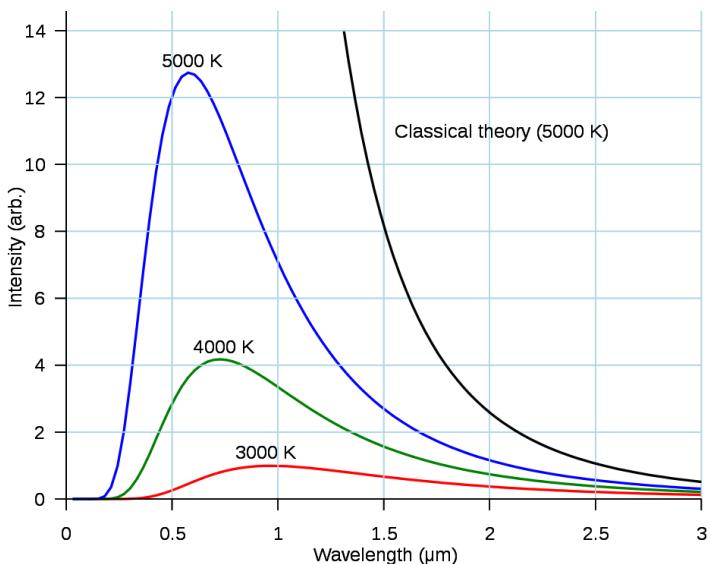
⁸⁰ http://en.wikipedia.org/wiki/Atmosphere_of_Earth

⁸¹ http://en.wikipedia.org/wiki/Black_body

The Planck function *spectral or specific brightness* may also be expressed as $B_v(T)$ which is used with frequency differentials dv , or as $B_{\bar{v}}(T)$ which is used with wave number differentials $d\bar{v}$ (where $\bar{v} = 1/\lambda$).

Blackbody radiation is a continuous spectrum with a peak emission intensity. The hotter a blackbody, the more the peak is shifted toward the higher energies (blue, UV, X-ray, etc.) The value of the Planck function for a particular blackbody and wavelength (or frequency) is always greater when T is greater, regardless how much the peak shifts (IMA2 p. 69).

The graph to the right depicts arbitrary intensity proportional to the spectral radiance (Planck function) for 3 temperatures, as well as the incorrect graph derived for 5000 K from classical theory.⁸²



Wien's displacement law

Derived by Wilhelm Wien (1864 - 1928) in 1893.⁸³ The wavelength λ_{\max} at maximum or peak blackbody spectral brightness or intensity $B_{\lambda}(T)$ (expressed as $\text{W m}^{-2} \text{nm}^{-1} \text{sr}^{-1}$) for a given equilibrium blackbody temperature is given by *Wien's displacement law*, where $b = 2.8978 \times 10^{-3} \text{ m K} = 2.8978 \times 10^6 \text{ nm}\cdot\text{K}$

$$\lambda_{\max} T = b$$

$$\lambda_{\max} = b / T$$

Hotter stars appear bluer, cooler stars appear yellower, orange, or redder. A very hot star having surface effective T of 42,000 K may have a Wien peak at 69 nm, far into the UV, and cool stars may have peaks in the infrared. The term "displacement" refers to the similarity among blackbody curves for differing temperatures, aside from the differing vertical scale and the displacement laterally of the peak.

Stefan-Boltzmann Equation

This relates Luminosity and Temperature. "Deduced by Jožef Stefan (1835–1893) in 1879, and derived from theoretical considerations, using thermodynamics, by Ludwig Boltzmann (1844–1906) in 1884."⁸⁴ Luminosity of a blackbody is given by

$$L = A\sigma T^4$$

where L = luminosity in W , T is temperature in Kelvins, A is area of the blackbody radiating surface, and σ is the Stefan-Boltzmann constant, a derived constant $= 5.670400 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$. Clearly, the luminosity is strongly dependent on and only on the temperature.

For a spherical star of radius R , this gives a relationship between luminosity and *effective temperature* T_e or T_{eff} :

$$L = 4\pi R^2 \sigma T_e^4 \quad \text{or}$$

$$T_e = T_{\text{eff}} = (L / 4\pi R^2 \sigma)^{1/4}$$

This relationship is used to express an effective temperature, even for objects that are far from having ideal blackbody spectral distribution. Note that it does not reference color index. The effective temperature is a property of the surface atmosphere or photosphere, not the interior temperatures.

⁸² http://en.wikipedia.org/wiki/Planck%27s_law

⁸³ http://en.wikipedia.org/wiki/Wien%27s_displacement_law

⁸⁴ http://en.wikipedia.org/wiki/Stefan%20%26amp;%20Boltzmann_law

For the Sun, Luminosity of Sun $L_{\odot} = 3.846 \times 10^{26}$ W, and Radius $R_{\odot} = 6.9599 \times 10^8$ m, giving

$$T_{\text{eff},\odot} = 5777 \text{ K}$$

Planck Blackbody Distribution Function (Planck Function)

Max Karl Ernst Ludwig Planck (1858 - 1947) derived his final 1900 formula, which included an empirical requirement that the allowable energy levels of “standing light waves” in a blackbody could only be integer multiples of a minimal energy $h\nu$ or hc/λ (i.e., $h\nu$, $2h\nu$, $3h\nu$, etc.), where h is a constant, $h = 6.62607 \times 10^{-34}$ J Hz⁻¹. (Eventually it was determined that these energy increments represent individual photons.)

The Planck function expresses the spectral brightness or spectral radiance, given as $B_{\lambda}(T)$, $B_{\nu}(T)$, or $B_{\bar{\nu}}(T)$, thus as functions of wavelength, frequency, or wave number. $B_{\lambda}(T)$ has units of $\text{W m}^{-2} \text{ nm}^{-1} \text{ sr}^{-1}$ (as discussed above), and is given by:

$$B_{\lambda}(T) = \frac{2hc^2/\lambda^5}{e^{\lambda/kT} - 1}$$

To compute radiant energy per unit time, one must integrate over the range of wavelengths, the area of the blackbody, and the solid angle into which emission is being calculated.

Integrating this formula over the area of an emitting star’s surface gives the *monochromatic luminosity* $L_{\lambda} d\lambda$ of the star

The *monochromatic flux* of a star (joules of starlight energy received per unit area with wavelength between λ and $\lambda + d\lambda$, and neglecting absorption or scattering) is given by $F_{\lambda} d\lambda$. (IMA2 p. 75)

Color and Color Index

Stars are viewed through special wide band filters with resulting *apparent color magnitudes*. The *Johnson-Morgan UBV stellar photometric system* (see for example their 1953 paper) uses filters of these color ranges (IMA2 p. 75):

U = Ultraviolet apparent magnitude m_U , with filter centered at 365 nm and effective bandwidth 68 nm
B = Blue apparent magnitude m_B , with filter centered at 440 nm and effective bandwidth 98 nm
V = Visual apparent magnitude m_V , with filter centered at 550 nm and effective bandwidth 89 nm.

These filters are extended in the Johnson-Cousins-Bessell system to UBVRI, adding:

R = Red apparent magnitude m_R , with filter centered at ~640 nm
I = InfraRed apparent magnitude m_I , with filter centered at ~798 nm

Additional filters further in the infrared have been used in the 2MASS all sky survey: J (1.25 μm), H (1.65 μm), and Ks (2.17 μm)⁸⁵—these aid in the search for cool brown dwarfs. L, M, and N filters have also been used.

The Sloan Digital Sky Survey SDSS uses u', g', r', i', z' (“ugriz”) filters with effective wavelengths of 360 (UV), 464 (green), 612 (red), 744 (IR), and 890 (IR) nm, respectively. There are a variety of additional systems and bands sampled. The filters used by HST’s WFPC2 have been previously mentioned.

Absolute color magnitudes M_U , M_B , and M_V may also be calculated if the distance is known.

Commonly used color indices of a star are:

$$\begin{aligned} B - V &= M_B - M_V \\ U - B &= M_U - M_B \end{aligned}$$

⁸⁵ <http://pegasus.phast.umass.edu/overview/overview.html>

These indices are used to quantitate color. The two quantities must be the same type of magnitude, either apparent magnitudes or absolute magnitudes, yielding equivalent results.

Stars with smaller B – V are bluer because lower or more negative magnitudes represent brighter sources. Stars with larger B – V are redder because higher or more positive magnitudes represent dimmer sources.

The color index of a star is a measure solely of the temperature of a model blackbody star and does not depend on stellar radius or distance (*IMA2* p. 77)

The magnitude of Vega has been used as the “zeropoint” calibration for visual and other color star magnitudes (see *IMA2* p. 76 and earlier discussion of zero point in this document). For each color filter, a constant is added to the flux measured to give the color magnitude, such that $U_{\text{vega}} = 0$. For example,

$$U = -2.5 \log_{10} \left(\int_0^{\infty} F_{\lambda} S_U d\lambda \right) + C_U$$

where S_U is a sensitivity of detection fraction for the instrument and C_U yields a U of approximately 0 for Alpha Lyrae (Vega). This does not imply that the actual brightnesses in U, B, and V bands for Vega are identical. A value $V_{\text{vega}} = 0.03$ has been chosen for the V band. Other systems of magnitude calibration are also used which are not tied to a particular star.

The bolometric magnitude is calibrated with

$$m_{\text{bol}} = -2.5 \log_{10} \left(\int_0^{\infty} F_{\lambda} d\lambda \right) + C_{\text{bol}}$$

The *Bolometric Correction BC* $\equiv m_{\text{bol}} - V$. Thus $m_{\text{bol}} = V + BC$. BC is therefore a constant to be added to V to get m_{bol} . C_{bol} was originally chosen to keep all values of BC negative, though some supergiant stars have positive values.

The difference of two magnitudes B_1 and B_2 comparing two stars is $-2.5 \log_{10} (N_1/N_2)$, where N_1 and N_2 are the total number of [photons] detected from a source above the background sky.⁸⁶

A graph of $U - B$ versus $B - V$ for a population of stars is not a simple straight line due to deviation of these stars from ideal blackbodies.

Note: There is some confusion when comparing blue-band or other band luminosity of an object to luminosity of the Sun. It is my understanding according to ZXI that unless otherwise stated, comparisons or ratios are made comparing:

$$L_{\text{object}}^{\text{Blue}} \text{ to } L_{\odot}^{\text{bol}}$$

rather than comparing

$$L_{\text{object}}^{\text{Blue}} \text{ to } L_{\odot}^{\text{Blue}}$$

AB Magnitude System

In the AB magnitude system,⁸⁷ AB magnitude is defined such that for any bandpass or filter being considered, the magnitude zero-point corresponds to a flux density of 3631 Jy. (1 Jy = 1 Jansky = $10^{-26} \text{ W m}^{-2} \text{ Hz}^{-1}$ = $10^{-23} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ Hz}^{-1}$). For monochromatic flux f measured in $\text{erg s}^{-1} \text{ cm}^{-2} \text{ Hz}^{-1}$

⁸⁶ Michael S. Bessell. *Annu. Rev. Astron. Astrophys.* 2005. 43:293–336, p. 296
Page 26 of 249

$$m(AB) = -2.5 \log(f) - 48.60$$

The zero point of the AB system uses a flux of 3631 Jy. Conversions with the Johnson system are given in the cited source. The AB system has been adopted by the Sloan Digital Sky Survey SDSS and the Galaxy Evolution Explorer (GALEX).

⁸⁷ AM Magnitude System:

- http://en.wikipedia.org/wiki/AB_magnitude
- <http://www.astro.utoronto.ca/~patton/astro/mags.html#conversions>

The Special Theory of Relativity (Chapter 4, omitted)

This chapter and topic has not been included in my astronomy courses to date. However, see my summary from a Physics 311 course [here](#).⁸⁸

Quantum Properties of Light and Matter; Spectroscopy (Chapter 5)

The quantum particle nature of light was postulated by Albert Einstein (1879 - 1955) in his 1905 analysis of the photoelectric effect. Each metal exhibits a characteristic cutoff frequency v_c whereby electrons are emitted only if the frequency of incident light $v > v_c$. Remarkably, this is independent of brightness. Einstein assumed photons exist with

$$E = hv = hc/\lambda$$

A 100W bulb emits roughly 2×10^{20} photons per second. Metals bind electrons, and a minimum binding energy must be provided to overcome this binding energy.

Spectroscopy

Spectral Lines

In the early 19C, Joseph von Fraunhofer (1787 - 1826) catalogued dark absorption lines in the spectrum of the Sun, and identified one as arising from sodium.

Gustav Kirchhoff (1824 - 1887) stated his laws of spectroscopy:

- 1) A hot dense gas or hot solid object produces a *continuous* spectrum [of light] with no absorption or emission lines. [This is a continuous blackbody spectrum.]
- 2) A hot, diffuse (thin) gas produces bright spectral lines (*emission lines*). [These are produced when the electron makes a downward transition to a lower orbit with concurrent emission of a photon.]
- 3) A cool, diffuse (thin) gas in front of a source of a continuous spectrum produces dark spectral lines (*absorption lines*) [These are produced when a photon is absorbed causing an upward transition to a higher orbit.]

Helium absorption lines were first identified in a solar spectrum in 1868, after which the element was found on Earth.

See below regarding naming of spectral lines.

Doppler Shifts

At non-relativistic slow speeds, approximate Doppler redshift z for the observed wavelength or frequency relative to the wavelength or frequency at rest at the emitter is given by (per IMA2 p. 112):

$$\begin{aligned} z &= (\lambda_{\text{obs}} - \lambda_{\text{emit}}) / \lambda_{\text{emit}} = \Delta\lambda / \lambda_{\text{emit}} = \lambda_{\text{obs}} / \lambda_{\text{emit}} - 1 = v_r/c, \quad \text{or} \\ z &= (f_{\text{emit}} - f_{\text{obs}}) / f_{\text{obs}} = \Delta f / f_{\text{obs}} = f_{\text{emit}} / f_{\text{obs}} - 1 = v_r/c \end{aligned}$$

Here, v_r is radial velocity of the emitter with respect to the observer.

Note the denominators, which differ “due to the fact that redshift causes wavelengths to increase and frequencies to decrease”⁸⁹. The Doppler shift z is positive if objects are moving apart (“redshift” to longer

⁸⁸ <http://www.mcgoodwin.net/pages/relativityprimer.html>

⁸⁹ <http://ned.ipac.caltech.edu/help/zdef.html>

wavelengths) or negative if objects are moving closer (“blueshift” to shorter wavelengths). See below for discussion of cosmological redshift, which can no longer ignore relativistic effects.

Light Bending and Spectrographs

Light can be *reflected* at a surface, in which the angle of incidence equals the angle of reflection.

It can be *refracted* (by passing from a medium with one index of refraction to one of a different index of refraction). Such refraction obeys *Snell's law*: $n_i \sin i = n_r \sin r$, where angle of incidence and of refraction are measured relative to the normal of the refracting surface.

Light may be *dispersed* when refracted in a prism or acted on by a diffraction grating. This *chromatic dispersion* occurs when the phase velocity and angle of refraction depends on the frequency.

Spectrographs are created with *diffraction gratings*, which are equivalent to multiple double slits in that they depend on constructive and destructive interference, and can be either transmissive or reflective. The smallest wavelength difference that can be resolved is given by λ/nN where n is the order (1, 2, 3...) of the spectrum and N is the number of slits. These spectra were originally captured on photographic plates, but are more commonly now detected by electronic detectors (CCDs).

Zeeman Effect

Splitting of spectral lines due to the *Zeeman Effect* (Pieter Zeeman 1865 - 1943, described in 1896) is used to measure stellar and other astronomical magnetic fields. In this phenomenon, the magnetic field resolves ordinarily unresolved (degenerate) electron configurations having the same energy into slightly different energies, which are apparent as split spectral lines. The degree of separation is proportional to the magnetic field except for high field strengths. This allows production of solar magnetograms and even measurement of weak magnetic fields in the ISM.

Causes of Spectral Line Broadening

Spectral Lines are *widened (broadened)* by several factors:⁹⁰

- “Natural Broadening” by the uncertainty principle—greater broadening ΔE occurs when the excited state is of shorter duration thus shortening Δt , as a result of the Uncertainty Principle ($\Delta E \Delta t \geq \hbar/2$).
- Doppler redshift and blueshift effects from rotation, thermal motion, and turbulence
- High pressure broadening and collisional broadening, due in part to frequent collisions again increasing uncertainty ΔE (seen for instance in white dwarf spectra) (IMA2 p. 270)
- Zeeman splitting, and
- Other inhomogeneities.

Light Interactions and Atomic Models

Compton Effect

Electrons inelastically scatter photons of wavelength λ_i , changing the wavelength in the process according to the Compton Effect formula:

$$\Delta\lambda = \lambda_f - \lambda_i = \frac{h}{m_e c} (1 - \cos \theta)$$

⁹⁰ IMA2 p. 268 and http://en.wikipedia.org/wiki/Spectral_line

where θ is the scattering angle of the photon. The electron is scattered into an angle φ . Clearly, the photon carries momentum even though it is massless. Arthur Compton discovered this in 1923. The effect is important because it demonstrates that light cannot be explained purely as a wave phenomenon.

Bohr Hydrogen Atom Model

Niels Bohr 1885 - 1962) developed his semiclassical model of the hydrogen atom (1913), which envisions electrons in circular orbits and yields a principle quantum number $n = 1, 2, 3, \dots$. The allowed energy levels are $-13.6 \text{ eV} * 1/n^2$. The values of n determine the energy level or “electron shell”.

Complete ionization of the electron from the ground state requires an input of 13.6 eV into the bound atom. The various hydrogen emission series are given by the Rydberg formula:

$$\frac{1}{\lambda} = \text{constants} \left(\frac{1}{n_{\text{Low}}^2} - \frac{1}{n_{\text{High}}^2} \right)$$

For instance, the Balmer series important in astronomy represents transitions to or from $n=2$ (which is at +10.19 eV relative to the ground level). The Lyman series arises from transitions to/from $n=1$ (which is at 0 eV, the ground level). The Paschen series is to/from $n=3$ (12.07 eV), Brackett is to/from $n=4$, and Pfund is to/from $n=5$.

A photon emitted from a transition to a lower energy level has energy equal to the difference in energy levels $\Delta E = h\nu = hc/\lambda f$. Similarly, when a photon is absorbed, the energy of the atom is raised by $\Delta E = h\nu = hc/\lambda$.

The Bohr model must be refined by fully quantum mechanical considerations with “fuzzy” orbitals (probability distributions) and additional quantum numbers.

Series of Spectral Lines

In astronomy, it is common to designate free atoms (not in molecular form) as

Neutral atom	H I, He I, C I, etc. (sometimes written HI, HeI, CI, etc., but this is less clear)
Singly ionized ions	H II (for H ⁺), He II (for He ⁺), O III (for O ⁺), etc.
Doubly ionized ions	He III (for He ⁺⁺), O IIII (for O ⁺⁺), etc.

(Note that H₂ (the diatomic hydrogen molecule) is not the same as H II. In addition, the cryogenic states of He, termed helium I and helium II (whether as isotopes ³He or ⁴He), is not the same as astronomical He I or He II.)

For unionized atomic hydrogen (termed H I), the commonly encountered spectral lines include those of the Balmer series, named according to the following pattern

- H α = 656.281 nm (n = 3 \Rightarrow 2 transitions),
- H β = 486.134 nm (n = 4 \Rightarrow 2 transitions),
- H γ = 434.048 nm (n = 5 \Rightarrow 2 transitions), and
- H δ = 410.175 nm (n = 6 \Rightarrow 2 transitions)

The Lyman lines Ly- α , Ly- β etc. are in the ultraviolet band. The Paschen and higher n-series lines are in the infrared. These non-visible light lines are best studied in high-altitude or space-based astronomy, due to atmospheric absorption.

Quantum Mechanical Considerations

Heisenberg Uncertainty Principle and Quantum Tunneling

The uncertainty (or indeterminacy) principle expresses a fundamental limitation on the simultaneous knowability of certain paired physical quantities ("conjugate variables"), and was stated by Werner Heisenberg (1901 - 1976) in 1927, modified in 1930. The Uncertainty Principle specifies uncertainties ΔE etc. for energy E, time of determination t, position x, and momentum p (the term "uncertainty" was not precisely defined):

$$\Delta E \Delta t \geq h \quad \Delta x \Delta p \geq h$$

Heisenberg's principle was given more precisely by E. H. Kennard in 1927 in terms of standard deviations σ and $h\bar{=}\hbar=h/2\pi$:

$$\sigma_E \sigma_t \geq \hbar/2 \quad \sigma_x \sigma_p \geq \hbar/2$$

but he was not describing simultaneous measurements. There remains considerable controversy regarding interpretation of the Uncertainty Principle.⁹¹

The uncertainties in position allow particles or photons to cross ("tunnel through") barriers in which classically they cannot exist. Tunneling⁹² through a potential barrier is more likely when the barrier is no more than a few wavelengths wide.

Wave-Particle Duality

The wavelike aspects of particles (*wave-particle duality*) were proposed by Louis de Broglie (1892 – 1987) in 1927. As with light photons, particle frequency ν is given by

$$\nu = E/h$$

and particle wavelength λ is given by

$$\lambda = h/p$$

where p is the momentum. This hypothesis was subsequently confirmed in double slit experiments. For electrons,

$$\lambda = h/m_e v$$

A sample calculation for a free electron moving at 0.01c shows the wavelength to be 250 pm, which is similar to the diameter of an unbound hydrogen atom (100 pm).

Schrödinger Equation and the Quantum Mechanical Atom

The waves de Broglie postulated are considered to be waves of particle positional probability, which are given by $|\Psi|^2$. The Ψ function is the Schrödinger equation, developed in 1925 by Erwin Schrödinger (1887 - 1961). Destructive interference of two particles (locations where interfering particles are not found) is seen where $|\Psi_1 + \Psi_2|^2 = 0$.

The Quantum Mechanics (QM) model of the atom introduces three additional quantum numbers. The angular momentum vector for the atom \mathbf{L} is determined by l (the azimuthal or angular momentum quantum number), which can take on values of $l = 0, 1, 2, \dots, n-1$. The quantum number l corresponds to the "subshell" orbitals, which have been named by the l values: s ($l=0$), p ($l=1$), d ($l=2$), f ($l=3$), g ($l=4$), h ($l=5$), etc. The magnitude of \mathbf{L} can take on only a set quantized levels determined by l (IMA2 p. 133):

$$L = \sqrt{l(l+1)}\hbar$$

The quantum number m_l (magnetic quantum number) determines possible values of the angular momentum z-component L_z . It specifies the specific orbital or "cloud" within the subshell, and can only take on values limited by l. (The following formula contains l, not the digit 1): $-l \leq m_l \leq l$

⁹¹ http://en.wikipedia.org/wiki/Uncertainty_Principle

⁹² http://en.wikipedia.org/wiki/Quantum_tunnelling

The quantum numbers l and m_l have no effect on the energy of a free isolated hydrogen atom.

However, if a magnetic field is present giving a preferred direction, the normally *degenerate* configurations (which all have the same energy and therefore produce the same spectral lines) acquire slightly differing energy levels due to differing values of m_l . The result is that “split” spectral lines can arise from the differing m_l values. Typically, three split lines are produced, not all of which are evident from a particular viewing angle (due to polarization effects). Thus, the line with unshifted frequency ν_0 corresponding to $m_l = 0$ may not be visible in a sunspot when the magnetic field is directed parallel to the line of sight, and only shifted lines are seen. As previously mentioned, this splitting is termed the Zeeman Effect. Even the ISM has a faint magnetic field, as low as 2×10^{-10} T, but this can be detected by measuring the Zeeman line shifting of hydrogen emissions.

A fourth quantum number m_s determines the orientation of the spin angular momentum of the electron. The overall magnitude of the spin angular momentum vector \mathbf{S} is fixed for a particular particle, and S is $(\frac{\sqrt{3}}{2})\hbar$ for the electron. The allowed values for the magnitude of the electron spin z-component S_z (where z direction is determined by the direction of the prevailing magnetic field) are $S_z = m_s\hbar = (\pm 1/2)\hbar$. Thus the quantum number m_s may take on values of $\pm 1/2$, depending on whether S_z is “spin up” or “spin down”. If the orientation of the electron spin angular momentum vector \mathbf{S}_z is up or *parallel* to the z-component of the proton’s angular momentum vector \mathbf{L}_z , it is at slightly higher energy (less unbound), compared to the *antiparallel* orientation which is slightly lower energy.

This leads to an additional potential cause of spectral line splitting, due to this “anomalous Zeeman Effect”. Such *hyperfine splitting* of the ground state is the source of the 1420.406 MHz 21 cm H line.

Note that for the H ground state, $n=1$, $l=0$, $m_l = 0$, and $m_s = \pm 1/2$. This state may be represented⁹³ by symbols H(1s¹) or the Russell-Saunders atomic term symbol ²S_{1/2}.

Pauli Exclusion Principle

Wolfgang Pauli (1900 - 1958) deduced in 1924-5 that no two electrons can occupy exactly the same quantum state, defined by the four quantum numbers (see above). This principle was derived by Paul Dirac (1902 - 1984) using relativistic wave equations. He showed further that this principle applies to all *fermions*. These are protons, neutrons, quarks, and all leptons including electrons, muons, taus, and neutrinos, etc.

Fermions all have spins which are half-integer multiples of \hbar , namely $\hbar/2$, $3\hbar/2$, $5\hbar/2$, etc.⁹⁴ (In contrast, *Bosons* have spins that are integer multiples of \hbar , namely 0, \hbar , $2\hbar$, $3\hbar$, etc., and they do not obey the Pauli Exclusion Principle.) See elsewhere in this document regarding degeneracy pressure.

Forbidden Transitions

Larger atoms have many possible energy levels and complex spectra result. Complex *selection rules* lead to *forbidden* vs. *allowed* transitions. Forbidden transitions may still occur but with much lower probability. For instance, the 21-cm line for Hydrogen is a forbidden transition, having a probability⁹⁵ of only $2.9 \times 10^{-15} \text{ s}^{-1}$. This 1420 MHz radiation arises from the transition between the two levels, F=1 and F=0, of the hydrogen ground state, which are slightly split by the interaction between the electron spin and the nuclear (proton) spin. The parallel (aligned) state of spin is slightly higher in energy (thus less loosely bound) than the antiparallel (anti-aligned) state.

In astronomy, forbidden line transitions are notated by placing square brackets around the atomic or molecular species in question, e.g., [O III] or [S II].

⁹³ Atomic Symbols and Russell-Saunders Term Symbol:

- <http://www.chem.ufl.edu/~itl/4412/lectures/ATermSym.html>
- http://en.wikipedia.org/wiki/Term_symbol

⁹⁴ <http://en.wikipedia.org/wiki/Fermion>

⁹⁵ http://en.wikipedia.org/wiki/Hydrogen_line

Telescopes (Chapter 6)

This is a limited survey of Chapter 6.

These topics are not summarized: *Refraction, Reflection, Snell's Law of Refraction, Focal Plane, Spherical Aberration, Chromatic Aberration* (with refractors), *Coma, Astigmatism, Field Curvature, Distortion, Magnification, Refracting and Reflecting Telescope Designs, Telescope Mounts, Large-Aperture Telescopes, Adaptive Optics, CCDs, Radio Telescopes, IR, UV, X-ray and Gamma ray telescopes, and All-Sky Surveys and Virtual Observatories.*

Viewing Conditions and Site Selection

Astronomical viewing conditions (*seeing conditions*) improve with reduction of interposed atmospheric turbulence (which causes "twinkling", water vapor, other absorbing gases, dust, etc. The HST Hubble Space Telescope addresses these needs, as will the James Webb Space Telescope JWST.

Ground based observatories attempt to locate at high altitude in dry locations: on Mauna Kea, these include Keck I and II, Gemini Northern, the Canada-France-Hawaii CFHT, Subaru, Univ. Hawaii, NASA Infrared IRTF, United Kingdom Infrared UKIRT, Caltech Submillimeter Obs. CSO, James Clerk Maxwell JCMT, the Submillimeter Array (SMA) and the Very Long Baseline Array VLBA. Other preferred locations for current and future instruments include the Atacama Desert, Chile and other Andean locales.

Focal Point and Focal Length f

Focal length is the distance between the center of a refracting objective lens or curved reflective mirror surface and its focus (focal point). The focal point is the point at which incoming light rays parallel to the optical axis are focused by the lens. Focal length is a measure of how strongly the lens converges or diverges light.

Plate Scale

For astronomical objects, the *plate scale* $\equiv d\theta/dy = 1/f$ is defined for each telescope and relates the angular separation $d\theta$ of the celestial sphere of an observed object to its linear separation dy at the focal plane of the detector. θ is in radians, f is the focal length, y and f are in identical units. This is the *IMA2* definition p. 144.

Plate scale may also be defined as the number of arcseconds that each pixel of the CCD detector spans (subtends).⁹⁶ In this case:

$$\text{plate scale (arcsec/pixel)} = s / d \\ \text{(where } s \text{ is individual pixel size and } d \text{ is overall CCD detector size.)}$$

Then

$$\text{plate scale (arcsec/pixel)} = s/d = s/(f/206265) = 206265 s/f \\ \text{where pixel size } s \text{ and focal length } f \text{ are in the same length units.}$$

Magnification

For traditional refracting telescopes, when the objective and eyepiece lenses are separated by the sum of their focal lengths, angular magnification of the virtual image is the ratio $f_{\text{obj}}/f_{\text{eye}}$ (the telescope objective focal length) to f_{eye} (the eyepiece lens focal length, *IMA2* p. 154, 175). This formula also applies to reflectors, but magnification is less useful especially for CCD detector systems, and plate scale is presumably a better index.

Resolution

Is based on an analysis of Airy disk interference for point sources. It specifies that two point sources are *unresolved* if the central maximum of one point's Airy pattern falls inside the location of the first minimum

⁹⁶ http://www.handsonuniverse.org/for_teachers/course/teacher_course/size2/size23.html

(circle of lowest intensity) of the other Airy pattern. This requirement leads to the *Rayleigh Criterion* for angular resolution:

$$\theta_{\min} = 1.22 \frac{\lambda}{D}$$

where θ is the minimum resolvable angle in radians of the object as viewed under ideal seeing conditions without a telescope, λ is the wavelength, and D is the telescope aperture diameter in the same units. Clearly, resolution improves with shorter wavelengths and wider apertures, all other factors being unchanged.

Focal Length and Focal Ratio (f-ratio)

The *focal ratio* F is given by f/D (where D is the aperture diameter). It also is called the f-ratio, f-stop, the f/number (written preceded by "f/", such as "f/2.8"), or the relative aperture. It is a unitless quantity as it is a ratio of identical length units.

Image Brightness

The image intensity of light (in $\text{W m}^{-2} \text{ nm}^{-1} \text{ sr}^{-1}$) from astronomical sources typically falls off only from attenuation and scattering, and otherwise the image intensity is equal to the object intensity (a somewhat surprising result of the geometric relationships and definitions, *IMA2* p. 153).

However, telescope light gathering effectiveness is expressed by *illumination J*, the amount of light energy (or number of photons) focused per unit time per unit area of the resolved image at the detector. The illumination J is proportional to $1/F^2$ or $(D/f)^2$, that is the inverse square of the f-ratio. If illumination is reduced in half, twice the time is required to gather the same number of photons. Comparing two telescope systems having f-ratios of 2 and 4, the one with $F=2$ has 4 times the illumination at the detector as the one with $F=4$.

Effect of Modification of Telescope Design Parameters

- Increasing telescope aperture D (keeping focal length f unchanged) increases illumination J and improves resolution (decreasing θ_{\min}).
- Increasing focal length f (keeping D unchanged) reduces illumination J but increases image size on the detector (spans more pixels).
- Increasing telescope aperture D (while keeping the focal ratio F unchanged, therefore increasing f proportionately) keeps illumination J constant, but improves resolution (lower θ_{\min}).
- Increasing focal length increases telescope support mass and complexity.
- Increasing aperture D increases mirror weight and potential for distortion, necessitates greater mirror support.

Binary Systems and Stellar Parameters (Chapter 7)

Binary star systems are widespread—more than 1/2 of all stars are in *multiple systems*, in which two or more stars are gravitationally bound and orbit about a common *center of mass* COM or *barycenter*). Some binary systems are useful for determining star masses and sizes. When differing in luminosity, etc., the more luminous is often termed the *primary* star, the other a *companion* star. Binary systems may be classified as follows:

- **[Optical Double]**: Not a binary system, simply two stars juxtaposed near a line of sight, but not gravitationally bound.]
- **Visual Binary**: Two stars optically resolvable due to adequate angular separation.
- **Astrometric binary**: A true binary where the primary is much brighter, and the unseen companion is detected only by oscillatory motion induced in the *position* of the primary star about the COM. This method is also used to detect extrasolar planets.
- **Eclipsing binary**: a binary with orbital plane approximately parallel to the line of sight, and for which the transit of one star in front of the other is detected by means of characteristic changes in the *light curve*. The effective temperatures and radii may often be inferred.
- **Spectrum binary**: a binary detectable by means of two superimposed independent discernible spectra. Separate blackbody-like spectra are usually detectable only when the cooler star is much larger in size (such as R Aquarii).
- **Spectroscopic binary**: If the stars have comparable size and luminosity, Doppler effects may cause detectable changes in the spectral lines of both stars (a *double-line spectroscopic binary*). When one star's lines are redshifted the other's will appear blueshifted due to opposing velocities in the orbital plane about the COM (see demo⁹⁷). If one star is much more luminous, the lines of the companion may be obscured but its presence may be detected by periodic fluctuation in the primary's lines.

Some binaries may satisfy more than one type named above, e.g., an eclipsing spectroscopic binary.

Binary systems may allow star **mass determination** in certain circumstances: visual binaries close enough to allow parallax distance determination, visual binaries with radial velocities available for full orbits; and eclipsing double-line spectroscopic binaries.

Many of the same kinds of techniques described for binary systems are now also used in the search for extrasolar planets.

Mass Determination Using Visual Binaries

To compute mass ratios: Two stars resolvable by adequate angular separation allow determination of the orientation of the orbits and the COM. The ratio of the masses is given from the semimajor axes if they are known:

$$\frac{m_1}{m_2} = \frac{\alpha_2}{\alpha_1}$$

where m_1 and m_2 are the stellar masses and α_1 and α_2 are the respective angular size of the semimajor axes of the two elliptical orbits about the COM. Thus the mass ratio can be computed even if the distance is unknown (IMA2 p. 184). Within limits, the orbital inclination plays no role in computing this ratio.

To compute individual masses: If distance d is known (by whatever means), the linear separation of the stars can be computed. Here the orbital inclination i , measured wrt a plane perpendicular to the line of sight, must be estimated. A relatively simple formula applies when the orbital plane has an inclination wrt the plane of the sky but there is no tilt about the major axis or rotation of the ellipse about the line of sight (IMA2 p. 185). Kepler's third law as modified may then be used to provide the sum of the masses:

$$m_1 + m_2 = \frac{4\pi^2}{GP^2a^3} \left(\frac{d}{\cos i} \right)^3 \frac{((\alpha_1 + \alpha_2)\cos i)^3}{P^2} = \frac{4\pi^2}{GP^2a^3} \left(\frac{d}{\cos i} \right)^3 \tilde{\alpha}^3$$

⁹⁷ <http://www.astro.cornell.edu/academics/courses/astro101/java/binary/binary.htm>

provided $\alpha = \alpha_1 + \alpha_2$, the semimajor axis of the orbit of the reduced mass, the inclination i is known, and the special conditions on the tilt as mentioned apply. The values for the ratio and sum of the masses may then be solved for the individual masses m_1 and m_2 . The possibility of additional types of angulation of the orbital plane makes the computation more complex.

Solutions for mass and other binary stellar parameters are often made with computerized data fitting techniques, comparing a synthetically generated light curve based on postulated exact parameters to the actual light curve and adjusting the synthetic parameters until the best fit is obtained.

Even if the distance is not known, it may still be possible to determine individual masses using Kepler's 3rd law if detailed radial velocity data are available.

Eclipsing Spectroscopic Binaries

Double-line eclipsing spectroscopic binary systems are especially useful for computing star parameters such as their masses, radii, ratio of their fluxes, and ratio of their effective temperatures (*IMA2* p. 186).

Mass Determination

The radial velocities of the individual stars, which determine the spectral line shifts, are influenced by the inclination, the ellipticity, the proximity, the rotation about the line of sight, etc.

See this demo⁹⁸ of radial velocities to better understand the effects of these factors. (In the demo, a is the orbital separation in solar radii determining period, e is the eccentricity, i is the inclination angle as seen from Earth, w is the “node angle—the angle of the major axis as measured in the orbital plane”, probably the same as orientation of the periastron).

Close binaries tend to circularize their orbits over short relative time scales due to tidal interactions. Stars in circular orbits have constant speed along their orbits given by $v = 2\pi a_1/P$, where P is period and a_1 is the semi-major axis of star 1. The ratio of star masses is given as

$$\frac{m_1}{m_2} = \frac{v_{2r}/\sin i}{v_{1r}/\sin i} = \frac{v_{2r}}{v_{1r}}$$

where v_{1r} is the observed radial velocity for star 1. Thus the ratio of masses is independent of the inclination.

To obtain the sum of the masses, we must know the inclination i . *IMA2* p. 188 derives:

$$\frac{m_2^3}{(m_1 + m_2)^2} \sin^3 i = \frac{P}{2\pi G} v_{1r}^3$$

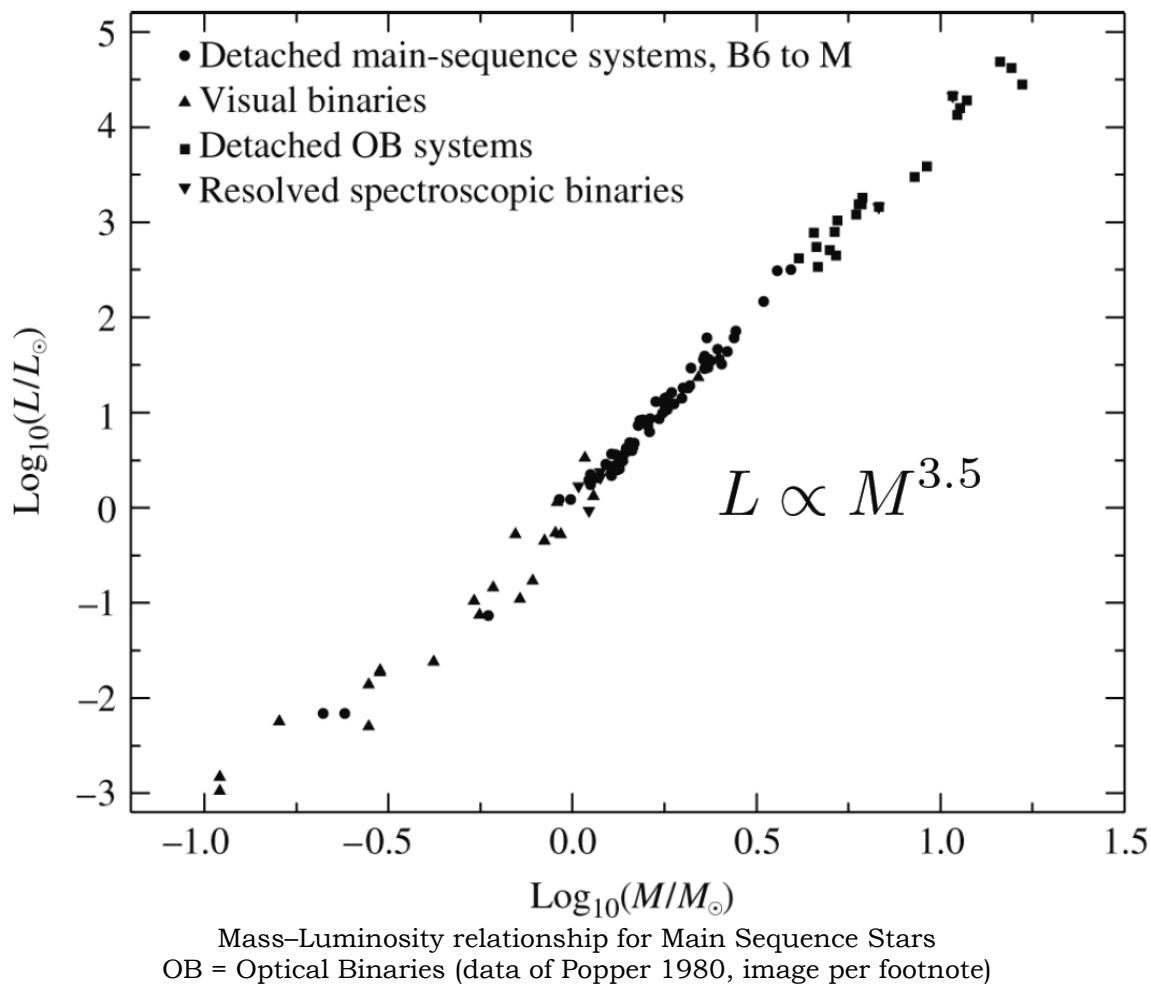
where the right hand side is the *mass function* requiring only the easily determined period and radial velocity of the primary star. Using this and the previous equation, masses can be determined if radial velocities of both and the inclination are obtainable. Even if only m_1 can be obtained, the mass function sets a lower limit for m_2 . As inclinations are often not known, these formulas are sometimes used for statistical studies giving mass estimates for a class of stars defined by luminosities and effective temperatures, etc.

In general, there is a mass-luminosity relationship for Main Sequence stars that has been derived using masses determined with visual and double line spectroscopic binaries. The empiric relationship is:

$$\frac{L}{L_\odot} = \left(\frac{M}{M_\odot}\right)^a$$

⁹⁸ <http://www.astro.cornell.edu/academics/courses/astro101/java/binary/binary.htm>
Page 36 of 249

where the exponent is found to be $3 < a < 4$, quoted as 3.5 for $2M_{\odot} < M < 20M_{\odot}$. Data showing this relationship is graphed below:⁹⁹



Inclination

Eclipsing binaries in most cases must have i (inclination) close to 90° relative to the plane of the sky. The specific pattern of light curves—specifically whether the minima are constant (aside from limb darkening effects) or curved—allows determination whether the geometry is partially or fully eclipsing and thereby improves the estimate of i .

Stellar Radii

The star radii may be determined from the amount of time between first contact t_a and the point of minimum light t_b as well as the star velocities. (*IMA2* p. 191):

$$r_s = (v/2)(t_b - t_a) \text{ for the smaller star, etc., where } v = v_s + v_l, \text{ the relative velocities of the two stars}$$

⁹⁹ image from <http://faculty.physics.tamu.edu/belyanin/astr314/lecture8.pdf> (A. Belyanin), similar to *IMA2* p. 189, adapted from data from Popper DMA, *Rev. Astr. Astrophys.* 18, 115 (1980)
Page 37 of 249

Ratio of Stellar Effective Temperatures T_{eff}

The ratio of the effective temperatures may be obtained from these light curves by assuming the stars are blackbody radiators (*IMA2* p. 192). The transit of a less luminous star in front of a more luminous star creates a deeper dip in the light curve than the opposite transit. In the textbook example, the ratio of effective temperatures may be derived from:

$$\frac{B_0 - B_p}{B_0 - B_s} = \left(\frac{T_s}{T_l}\right)^4$$

where B_0 is the brightness when both stars are fully seen, B_p is brightness of the primary minimum (when the smaller hotter star is eclipsed), and B_s is the brightness of the secondary minimum. T_s is the effective temperature of the smaller star.

The Classification of Stellar Spectra (Chapter 8)

Harvard Classification

Stellar classification by spectra began with Harvard's Edward Pickering (1846 - 1919) and Williamina P. Fleming in the 1890s using capital letters, as well as by Antonia Maury. Astronomer Annie Jump Cannon (1863 - 1941) gave us the final sequence of the first 7 letters O B A F G K M, and their subdivision into A0 - A9 etc., based on the strength of the hydrogen Balmer lines. Her observations on 20,000 stars were compiled into the *Henry Draper Catalogue*, from which stellar HD numbers derive. This taxonomy is termed the *Harvard Classification*,¹⁰⁰ and is based (as became clear only later) on a sequence of decreasing stellar surface temperatures. The currently expanded sequence of letters is **O B A F G K M R N S L T** etc., for which there are some amusing mnemonics.¹⁰¹ (The letter *P* in the Draper catalog referred to Planetary Nebula, and *Q* were "other spectra"—letters omitted currently.) O stars are the hottest (and therefore newest), with surface temperatures of $\geq 33,000$ K. L and T are very cool stars and brown dwarfs. Our Sun is a G2V star (V is the luminosity class, see below). The stars near the beginning (O B A) are still termed "*early-type*" and the stars near the end (K M) are "*late-type*", terminology reflecting now-discredited models of stellar evolution (*IMA2* p. 220).

The Harvard Spectral Classification¹⁰²

Data shown for M_{\odot} , R_{\odot} , and L_{\odot} apply to main sequence [luminosity class V] stars only

Class	Surface Temp K	Conventional color	Apparent color	Mass M_{\odot}	Radius R_{\odot}	Luminosity (Bolometric) L_{\odot}	Hydrogen Lines	% of Main Sequence Stars
O	$\geq 33,000$	blue	blue	≥ 16	≥ 6.6	$\geq 30,000$	Weak	~0.00003%
Hottest blue-white stars, few lines, Strong He II absorption (sometimes emission), He I lines becoming stronger.								
B	10,000–33,000	blue to blue white	blue white	2.1–16	1.8–6.6	25–30,000	Medium	0.13%
Hot blue-white, He I absorption lines strongest at B2, H I n=2 (Balmer) abs. lines becoming stronger.								
A	7,500–10,000	white	white to blue white	1.4–2.1	1.4–1.8	5–25	Strong	0.6%
White, Balmer abs. lines strongest at A0, becoming weaker later. Ca II abs. lines becoming stronger.								
F	6,000–7,500	yellowish white	white	1.04–1.4	1.15–1.4	1.5–5	Medium	3%
Yellow-white. Ca II lines continue to strengthen as Balmer lines continue to weaken. Neutral metal abs. lines (Fe I, Cr I)								

¹⁰⁰ http://en.wikipedia.org/wiki/Stellar_classification#Harvard_spectral_classification

¹⁰¹ OBAFGKMRNSLT mnemonic:

- Oh be a fine girl, kiss me
- Oh Brother, Astronomy Finally Gruesomely Killed Me Right Now (Slump)
- Oven Baked Ants, Fried Gently, Kept Moist, Retain Natural Succulence (Largely True)

¹⁰² Stellar Classification:

- http://en.wikipedia.org/wiki/Stellar_classification

- *IMA2* p. 205

- <faculty.physics.tamu.edu/belyanin/astr314/lecture8.pdf>

G	5,200–6,000	yellow	yellowish white	0.8–1.04	0.96–1.15	0.6–1.5	Weak	7.6%
Yellow. Sun is G2. Ca II lines continue becoming stronger. Fe I, other neutral metal lines becoming stronger.								
K	3,700–5,200	orange	yellow orange	0.45–0.8	0.7–0.96	0.08–0.6	Very weak	12.1%
Cool orange. Ca II H and K lines strongest at K0, becoming weaker later. Spectra dominated by metal abs. lines.								
M	≤ 3,700	red	orange red	≤ 0.45	≤ 0.7	≤ 0.08	Very weak	76.45%
L	Very cool, dark red (brown dwarfs). Brighter in Infrared than visible. Strong molecular absorption bands of metal hydrides (CrH, FeH), water, CO, and alkali metals (Na, K, Rb, Cs). TiO and VO weakening							
T	Coolest stars or brown dwarfs. Strong methane (CH_4), weakening CO bands.							

The letters R N and S were for peculiar stars (with high amounts of C, lanthanum oxide, or zirconium oxide, etc.) and are no longer used. There are several published catalogs of stellar spectra.¹⁰³

Other Libraries and Models of Stellar Spectra

In addition to the textbook (IMA2 p. 205), may sources graph, tabulate, and compare the physical characteristics of stars emitting the various spectra¹⁰⁴—I am yet to make a careful study of this huge body of information. ZXI commends the Robert L. Kurucz models of *Stellar Energy Distributions* (SEDs) [aka *Spectral Energy Distributions*]¹⁰⁵ as a means of understanding how these spectra arise. He states that the Kurucz models make use primarily of the effective temperature, surface gravity, and metallicity.

Statistical Mechanical Considerations

A fuller understanding of the source of these spectra required analysis not only of electron orbital probabilities and also the probability of ionization given by statistical mechanics.

The Maxwell-Boltzmann velocity distribution function (IMA2 p. 206 and *here*¹⁰⁶), the most probable speed of an atom $v_{\text{mp}} = (2kT/m)^{1/2}$, and the root mean square speed of an atom $v_{\text{rms}} = (2kT/m)^{1/3}$ are illustrated as follows for hot hydrogen atoms (graph does not yet consider the likelihood of ionization):

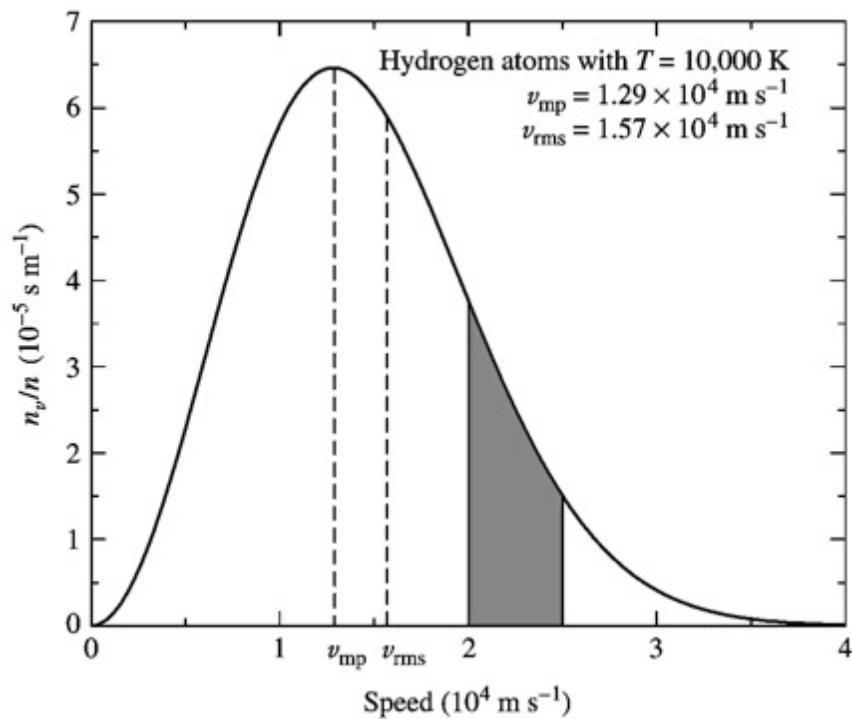
¹⁰³ • Jacoby, G. H., Hunter, D. A., & Christian, C. A., “A library of stellar spectra”, *Astrophysical Journal Supplement Series* (ISSN 0067-0049), vol. 56, Oct. 1984, p. 257-281, available online in PDF form.

• www.annualreviews.org/doi/pdf/10.1146/annurev.astro.42.053102.134017 for L and T stars & brown dwarfs
¹⁰⁴ http://en.wikipedia.org/wiki/Stellar_classification

¹⁰⁵ Sample Kurucz papers and models of Stellar Energy Distributions (SEDs):

- <http://www.mso.anu.edu.au/observing/2.3m/CASPIR/docs/manual/node228.html>
- Kurucz, R. L., “Model atmospheres for G, F, A, B, and O stars”, *Astrophysical Journal Supplement Series*, vol. 40, May 1979, p. 1-340.
- R. Buser and R.L.Kurucz. “A library of theoretical stellar flux spectra. I. Synthetic UBVRI photometry and the metallicity scale for F- for K-type stars” *A&A* 264, 557-591 (1992) [340 pages long!]

¹⁰⁶ http://en.wikipedia.org/wiki/Maxwell-Boltzmann_distribution



Maxwell-Boltzmann Probability Density Function for Hydrogen Atoms at $T = 10,000\text{K}$
 The fraction of hydrogen atoms in the gas having velocities between $2 \times 10^4 \text{ m s}^{-1}$ and $2.5 \times 10^4 \text{ m s}^{-1}$ is the shaded area under the curve between those two velocities divided by the entire area under the curve.¹⁰⁷

The Boltzmann Equation

is given as follows:

$$\frac{P(S_b)}{P(S_a)} = e^{-(E_b - E_a)/kT}$$

where $P(S_b)$ is the probability that a system is in state S_b , E_b is the energy of this state (based on the quantum numbers characterizing this state), T is the temperature in K, and k is the Boltzmann constant. This equation shows that higher energy orbitals are statistically less likely to be occupied than lower energy orbitals, all other factors being equal.

The important and recurring quantity

$$e^{-E/kT}$$

is called the *Boltzmann factor*. At room temperature (300 K), $kT = 1/40 \text{ eV}$.

States S_b , and S_a , with equal energies ($E_b = E_a$) but which are not identical (having identical quantum numbers) are termed *degenerate*. The presence of multiple degenerate states when present must be accounted for in the expression of the Boltzmann equation.

The final Boltzmann Equation applied to stars (for which probabilities are proportional to numbers of atoms), and including g factors allowing for degeneracy, and again not accounting for ionization, is given as follows (IMA2 p. 212):

$$\frac{N_b}{N_a} = \frac{g_b}{g_a} e^{-(E_b - E_a)/kT}$$

where N_b is the relative number of atoms in state S_b , E_b is the energy of this state (based on the quantum numbers characterizing this state), T is the temperature in K, and k is the Boltzmann constant.

¹⁰⁷ <http://faculty.physics.tamu.edu/belyanin/astr314/lecture8.pdf> after IMA2 p. 211.

If one attempts to use this equation to find at what temperature there are equal numbers of unionized H atoms in the ground state ($n=1$) and the next higher state ($n=2$), the answer computed is 85,400K. This is incorrect, because ionization has not been considered (see to follow).

Saha Equation

Developed in 1920 by astrophysicist Meghnad (or Megh nad) Saha (1894 - 1956), this gives the relative number of atoms in different states of ionization. Ionization energies needed to go from state i to next higher level of ionization $i + 1$ are represented by χ_i . The ratio of the number of atoms in ionization stage ($i + 1$) to the number in state (i) is given by:

$$\frac{N_{i+1}}{N_i} = \frac{2kT Z_{i+1}}{P_e Z_i} \left(\frac{2m_e kT}{h^2} \right)^{3/2} e^{-\chi_i/kT}$$

where m_e is the electron mass, P_e is the pressure of free electrons (which can often only be estimated), and the Z_i terms are partition functions incorporating degeneracy factors g_j :

$$Z = \sum_{j=1}^{\infty} g_j e^{-(E_j - E_1)/kT}$$

For stars, electron pressure in the atmospheres ranges from 0.1 to 100 N m⁻².

Computing the Strength of Balmer Lines versus Calcium Lines, etc.

In practice, the Boltzmann and Saha equations are combined to compute stellar spectra, for instance by computing the relative numbers of $n=2$ hydrogen atoms available to create Balmer lines. For example, Saha shows us that for H, 5% of atoms are ionized at 8300K, 50% are ionized at 9600K, and 95% are ionized at 11,300 K. This 3000K range (from 5% to 95% ionization) is termed the *partial ionization zone*, and is conveniently approximated for H as 10,000K. The Balmer lines (H α , H β , etc.) are seen maximally at 9250 K because at temperatures higher than 10,000K, H is nearly fully ionized so cannot produce Balmer lines. The population of H in the state $n=2$ peaks at 9,900 K (assuming $P_e = 20$ N m⁻²), and at higher temperatures there is a rapid falloff of $n=2$ H I atoms to provide Balmer lines. When other atoms are present, such as He, there are additional electrons to recombine with ionized H, so a higher temperature is required to achieve the same degree of H ionization.

After computing with Saha Equation the fraction on atoms which are in a particular state of ionization, one uses the Boltzmann Equation to obtain the number of atoms in a particular excitation state available to create spectral lines.

For the Sun with $T_e = 5777$ K, only a tiny fraction of H is in the $n=2$ H I state, and Balmer lines are inconspicuous, whereas Ca II H and K lines are strong (IMA2 p. 219).

Compositions of Stellar Atmospheres

Cecilia Payne (1900-1979) was the first to calculate the relative compositions of stellar atmospheres in her 1925 thesis.

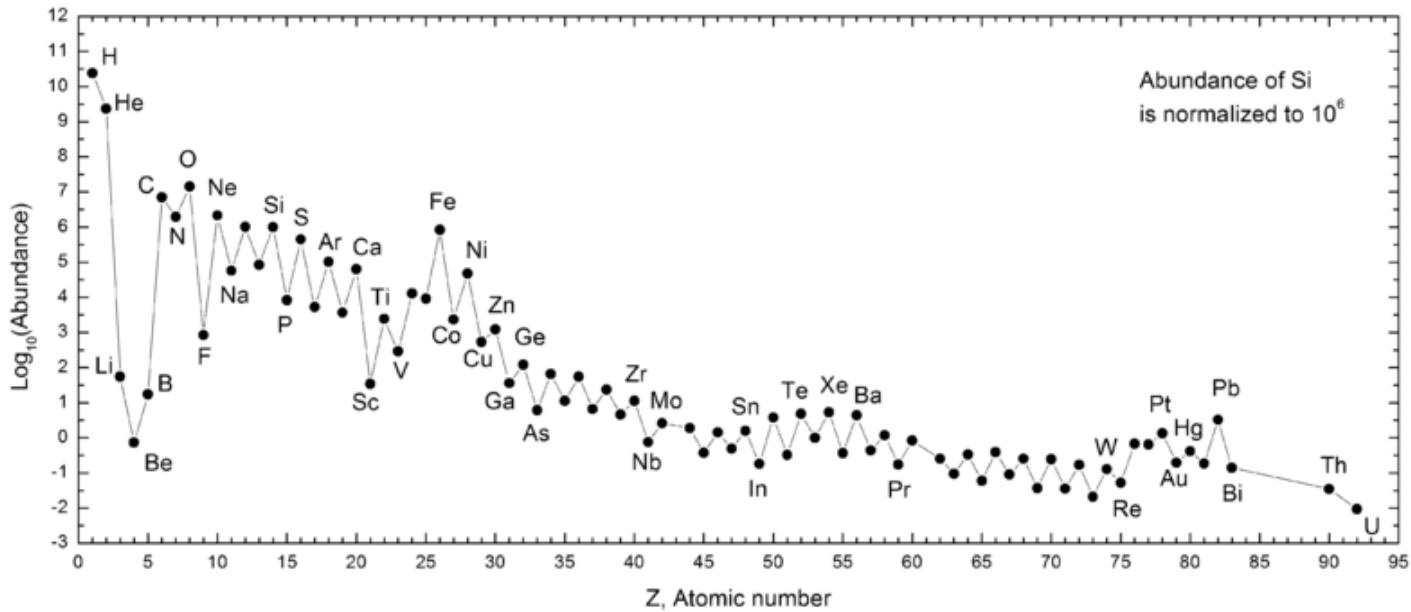
Here is a table of the 10 most common elements in the Sun (apparently based on solar spectra)¹⁰⁸:

¹⁰⁸ http://imagine.gsfc.nasa.gov/docs/ask_astro/answers/961112a.html See also IMA2 p. 277
Page 42 of 249

Element	Solar Abundance (% of total number of atoms)	Solar Abundance (% of total mass)
Hydrogen	91.2	71.0
Helium	8.7	27.1
Oxygen	0.078	0.97
Carbon	0.043	0.40
Nitrogen	0.0088	0.096
Silicon	0.0045	0.099
Magnesium	0.0038	0.076
Neon	0.0035	0.058
Iron	0.030	0.014
Sulfur	0.015	0.040

The nine most abundant cosmic nuclear species in descending order are: hydrogen-1, helium-4, oxygen-16, carbon-12, neon-20, nitrogen-14, magnesium-24, silicon-28, and iron-56 (IMA2 p. 315)

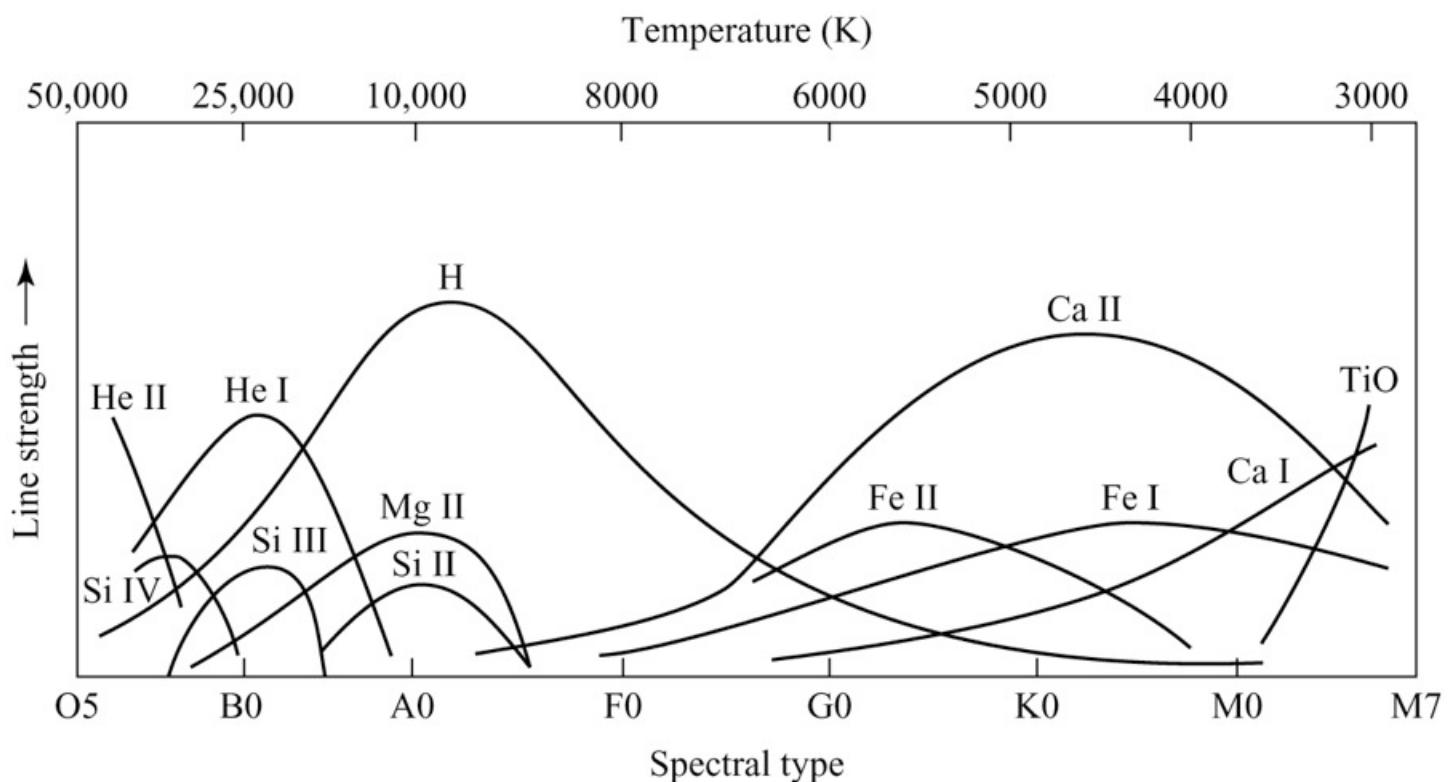
Here is a chart showing the **number abundances of the elements in the solar system** (not just the Sun's photosphere).¹⁰⁹ The textbook (IMA2 p. 541) shows a similarly plotted graph for the relative abundance of the Sun's photosphere. H and He are of course much more abundant there.



The following graph indicates the relative spectral line strengths of various elements and the TiO molecule in stellar atmospheres as a function of spectral type or temperature¹¹⁰:

¹⁰⁹ http://en.wikipedia.org/wiki/Abundance_of_the_chemical_elements

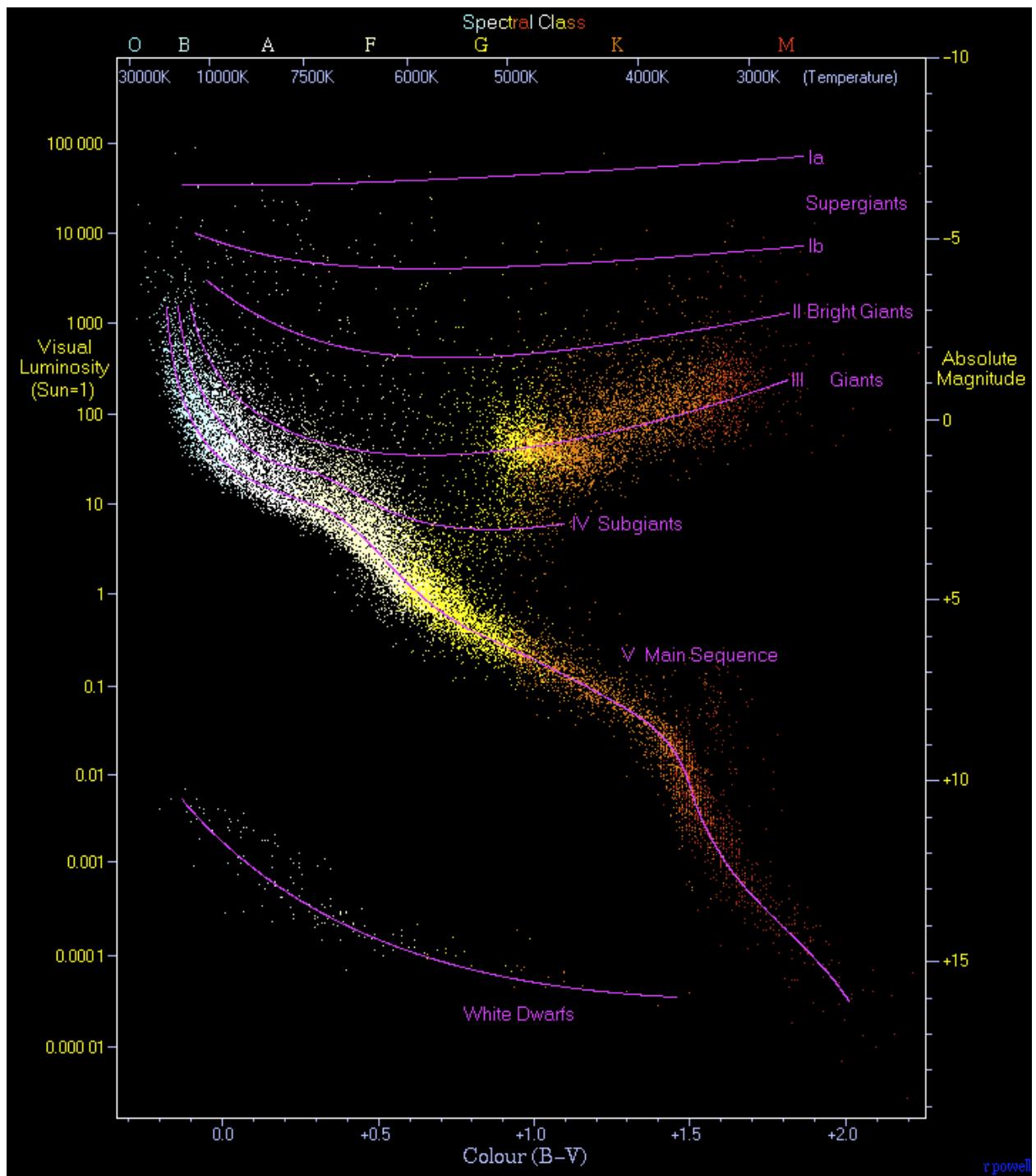
¹¹⁰ <http://faculty.physics.tamu.edu/belyanin/astr314/lecture8.pdf> adapting IMA2 p. 220



The Hertzsprung-Russell (H-R or HD) Diagram

The data was compiled in tabular form c. 1910 by Ejnar Hertzsprung (1873 - 1967) and Henry Norris Russell (1877 - 1957), initially as a means of studying stellar evolution. Russell published the information in graphical form in 1913. An excellent simulation of stellar evolution on a H-R diagram grid is given here¹¹¹. A modern version follows:

¹¹¹ <http://www.astro.ubc.ca/~scharein/a311/Sim/hr/HRdiagram.html> requires JAVA
Page 44 of 249



A modern Hertzsprung–Russell diagram¹¹²
with 22,000 stars plotted from the Hipparcos catalog
and 1,000 from the Gliese catalog of nearby stars.

The Sun, a G2V star, is located at absolute magnitude 4.8, luminosity 1, and B-V=+0.66

“The ordinary hydrogen-burning dwarf stars like the Sun are found in a band running from top-left to bottom-right called the Main Sequence. Giant stars form their own clump on the upper-right side of the diagram. Above them lie the much rarer bright giants and supergiants. At the lower-left is the band of white dwarfs—these are the dead cores of old stars which have no internal energy source and over billions of years slowly cool down towards the bottom-right of the diagram.

¹¹² Hertzsprung–Russell diagram:

- image and most text that follows it from <http://www.atlasoftheuniverse.com/hr.html> (Richard Powell)
- a few comments also from http://en.wikipedia.org/wiki/Hertzsprung%20%93Russell_diagram

Stellar Spectral Energy Distributions (SEDs) are a function of mass, chemical composition, and age. *Kurucz models* (1979) assume main sequence star SEDs are determined by effective temperature, $\log(g)$, and metallicity (expressed as [Fe/H]) (ZXI notes).

“The original [H-R] diagram displayed the *spectral type* of stars on the horizontal axis and the absolute magnitude on the vertical axis. The first quantity (i.e. spectral type) is difficult to plot as it is not a numerical quantity, and in modern versions of the chart it is replaced by the *B-V color index* of the stars.”¹¹³

After stars have consumed much of their H, they turn off the main sequence (a period called the Post-Main Sequence).

Morgan-Keenan Luminosity Classes and the M-K Classification System

In addition to the Harvard Classification of star spectral classes, stars are also classified into five main *luminosity classes*. Stars having the same spectral type (thus same effective temperature), were noted by Antonia Maury and Hertzsprung to vary in line width. This observation culminated in the 1943 publication of the *Atlas of Stellar Spectra (“MKK Atlas”)* by William W. Morgan (1906 - 1994), Phillip C. Keenan (1908 - 2000) and Edith Kellman (1911 - 2007), all of Yerkes Observatory. Their MKK atlas and classification adds a luminosity class (designated by Roman numerals I to V) to the Harvard classification spectral type. For stars of the same spectral type, narrower lines (lower Roman numeral M-K luminosity classes) are seen with stars of higher luminosity. (This is because they have thinner atmospheres and their atoms experience fewer atomic collisions, causing less collisional broadening of lines.)

These are the five luminosity classes (quoted from here¹¹⁴):

I Supergiants

Very massive and luminous stars near the end of their lives. They are subclassified as Ia or Ib, with Ia representing the brightest of these stars. These stars are very rare - 1 in a million stars is a supergiant. The nearest supergiant star is Canopus (F0Ib) 310 light years away. Some other examples are Betelgeuse (M2Ib), Antares (M1Ib) and Rigel (B8Ia).

II Bright Giants

Stars which have a luminosity between the giant and supergiant stars. Some examples are Sargas (F1II) and Alphard (K3II).

III Normal Giants

These are mainly low-mass stars at the end of their lives that have swelled to become a giant star. This category also includes some high mass stars evolving on their way to supergiant status. Some examples are Arcturus (K2III), Hadar (B1III) and Aldebaran (K5III).

IV Subgiants

Stars which have begun evolving to giant or supergiant status. Some examples are Alnair (B7IV) and Muphrid (G0IV). Note also Procyon which is entering this category and therefore is: F5IV-V.

V Dwarfs

All normal hydrogen-burning stars. Stars spend most of their lives in this category before evolving up the scale. Class O and B stars in this category are actually very bright and luminous and generally brighter than most Giant stars. Some examples are the Sun (G2V), Sirius (A1V), and Vega (A0V). “

The M-K classification has been extended (*IMA2* p. 225) to include subdwarf stars (VI or “sd” but not universally to white dwarfs (“D”).¹¹⁵ Other authors include a VII class for white dwarfs (aka “wD” or “WD”).

Other luminosity class extensions sometimes used include “0” [zero] and 1a-0 for hypergiants (e.g., Rho Cassiopeiae is a G2Ia0e yellow hypergiant. Extension “Iab” signifies intermediate between Ia and Ib). Final extension e means “emission lines present”.

Spectral *peculiarity* designations appended at the end, such as e, ep, k, m, n, p, s, var, v, !, etc., are summarized here.¹¹⁶

¹¹³ http://en.wikipedia.org/wiki/Hertzsprung%20%93Russell_diagram

¹¹⁴ <http://www.atlasoftheuniverse.com/hr.html>

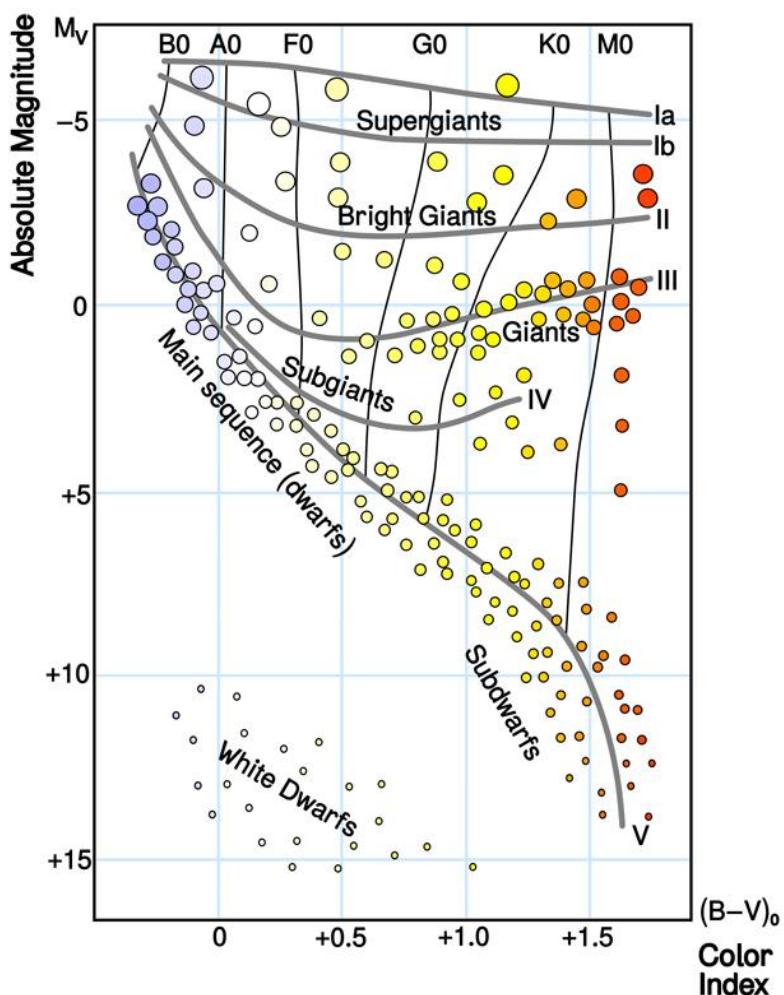
¹¹⁵ However, the table on *IMA2* p. 228 shows D white dwarfs listed as part of the M-K luminosity class system
Page 46 of 249

The radius of a star can be determined from its position in the H-R diagram by the Stefan-Boltzmann law. On a logarithmically plotted H-R diagram, stars of the same radii fall along diagonal lines that run roughly parallel to the main sequence.

The position of a main sequence star on the main sequence (luminosity class V) is determined primarily by its mass (see IMA2 App. G). The Main Sequence stars have a density similar to water. The Sun, a G2V star, has average $\rho_{\odot} = 1.4 \text{ g cm}^{-3}$. Early-type stars have lower average density: Sirius A, an A1V star, has average $\rho = 0.76 \text{ g cm}^{-3} = 0.54 \rho_{\odot}$. In contrast, the giant stars can have very low average density. Betelgeuse, a supergiant M2Iab star, has an average density of only $10^{-8} \rho_{\odot}$.

Using the M-K system, astronomers can locate a star on the H-R diagram, thereby estimate its absolute magnitude, and then use its apparent magnitude to estimate its distance using the distance modulus. This technique of estimating distance is loosely called by the misnomer *spectroscopic parallax*.

The adjacent graph¹¹⁷ shows the relationship between spectral classes such as F0, color index, luminosity classes (lines labeled with Roman numerals), and absolute visual magnitudes M_V . Note that when the horizontal axis is color index and the vertical axis is absolute magnitude, the luminosity classes and the boundaries separating the spectral classes are all curving lines. (In some diagrams, one is led to the mistaken impression that the spectral classes are demarcated by vertical lines.)



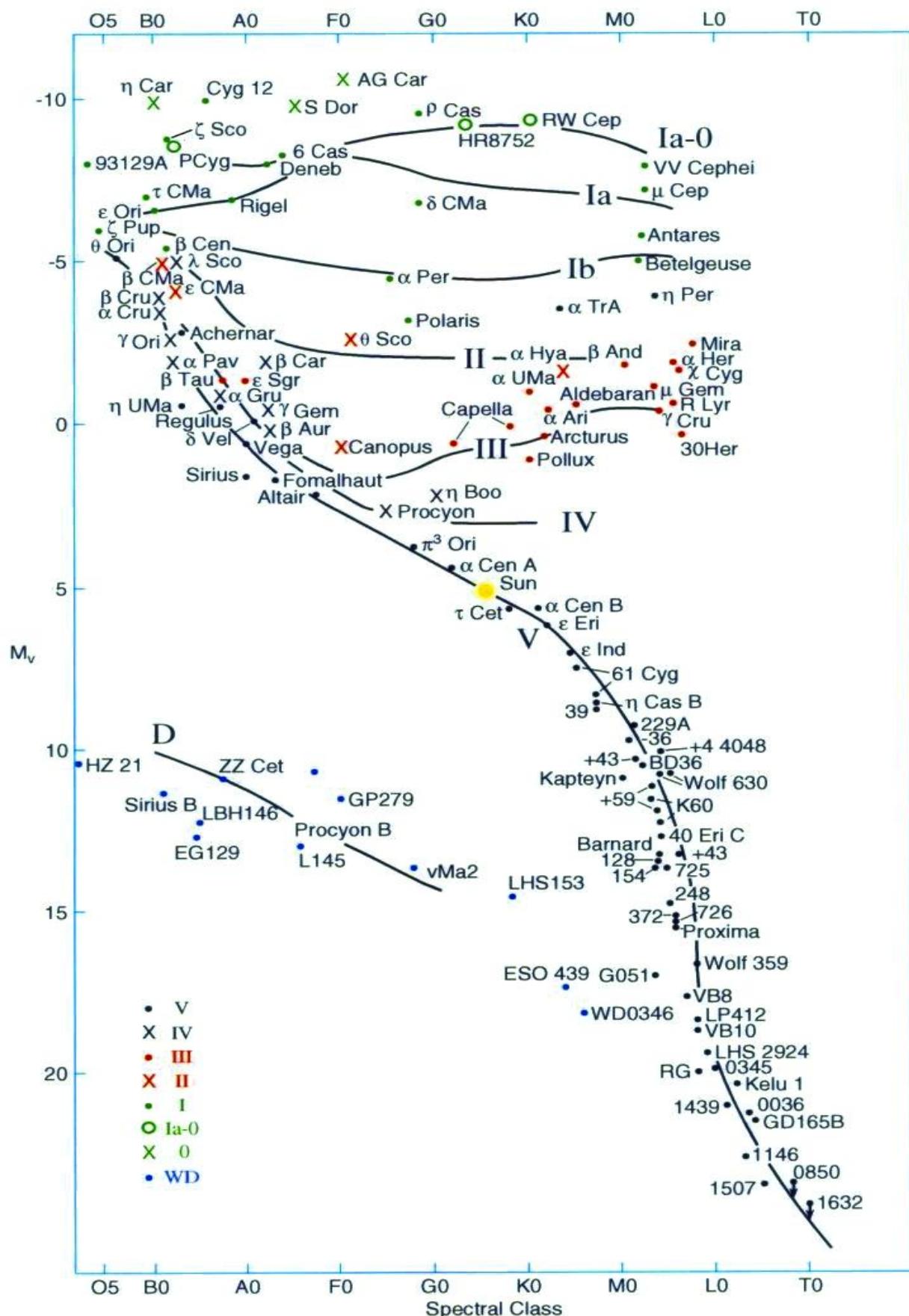
Another image of the H-R diagram, also including many non-main-sequence stars, follows:¹¹⁸

¹¹⁶ Stellar Spectral peculiarity designations:

- http://en.wikipedia.org/wiki/Stellar_classification#Spectral_peculiarities
- <http://www.skyandtelescope.com/howto/basics/3305876.html?showAll=y>

¹¹⁷ http://en.wikipedia.org/wiki/Stellar_classification

¹¹⁸ image (modified by MCM) fr. <http://stars.astro.illinois.edu/sow/hrd.html>, fr. JB Kaler, *The Cambridge Encyclopedia of Stars*, Cambridge U Press, Nov 2006



Stellar Atmospheres (Chapter 9, omitted)

This chapter has not yet been read in my astronomy courses and is not summarized here. Major topics discussed in this chapter are:

The Radiation Field

- Specific Intensity and Mean Intensity
- Specific Energy Density
- Specific Radiative Flux
- Radiation Pressure

Stellar Opacity

- Temperature and Local Thermodynamic Equilibrium
- Definition of Opacity and Monochromatic Opacity κ_λ (defined by $dI_\lambda/ds = -\kappa_\lambda I_\lambda \rho$)
- Optical Depth τ_λ
- Sources of Opacity
- Continuum Opacity and the H⁻ ion
- The Rosseland Mean Opacity $\bar{\kappa}$

Radiative Transfer

- Photon Emission Processes
- The Random Walk
- Limb Darkening (of a star at the edge)
- The Radiation Pressure Gradient

The Transfer Equation

- The Emission Coefficient j_λ
- The Special Case of Blackbody Radiation
- The Assumption of Plane-Parallel Atmosphere
- The Eddington Approximation
- Limb Darkening Revisited

The Profiles of Spectral Lines

- Equivalent Widths
- Processes that Broaden Spectral Lines
- The Voigt Profile (flux plotted against wavelength)
- The Curve of Growth
- Computer Modeling of Stellar Atmospheres

The Interiors of Stars (Chapter 10)

This chapter discusses stellar models and derives some of the needed formulas. It is only partially summarized here. Stars are “gravitationally confined thermonuclear reactors whose composition evolves as energy is lost to radiation and neutrinos.”¹¹⁹

Basic Equations

Hydrostatic Equilibrium

The relevant formula for a spherically symmetrical star or other body is derived (*IMA2* p. 287):

$$\frac{dP}{dr} = -G \frac{M_r \rho}{r^2} = -\rho g$$

where P is pressure, r is radius from the center of the star, M_r is the mass contained within a sphere of radius r, ρ is the local density, and g is the local acceleration of gravity = GM_r/r^2 . When applied using computer modeling, a central solar pressure of nearly 2.34×10^{16} N m⁻² is computed, equivalent to 2.3×10^{11} Earth atm.

The Equation of Mass Conservation

Again for a spherically symmetrical star (*IMA2* p. 288),

$$\frac{dM_r}{dr} = 4\pi r^2 \rho$$

Pressure Equation of State

A **state equation (equation of state EOS)** is a relationship between state variables—temperature, pressure, volume, entropy, enthalpy, internal energy, etc.—whose values depend only on the current state of the system and not the path by which the state was attained. The thermal gas pressure is given by

$$P_g = \frac{\rho k T}{\mu m_H}$$

where P_g is the local ideal pressure of the gas, ρ is local density, k is Boltzmann constant, T is temperature, and μ is the mean molecular weight of the gas.¹²⁰ (For stars with ionized gases, this is actually the mean particle weight, counting free electrons as well as ions as individual particles):

$$\mu \equiv \frac{\bar{m}}{m_H}$$

where \bar{m} is the average mass of a gas particle (including numbers of electrons in the denominator), and m_H is the mass of a hydrogen atom = 1.67×10^{-27} kg. Clearly, the mean molecular weight depends on the state of ionization. (*IMA2* p. 288)

The Average Kinetic Energy Per Particle

This is given by

¹¹⁹ Woosley SE, et al, “The evolution and explosion of massive stars”, *Rev. Mod. Phys.* 74, 1015–1071 (2002)

¹²⁰ http://www.vikdhillon.staff.shef.ac.uk/teaching/phy213/phy213_molecular.html

$$\frac{1}{2}mv^2 = \frac{3}{2}kT$$

which reflects the three coordinate directions or degrees of freedom (*IMA2* p. 294).

Radiation Pressure

The contribution to pressure attributable to photon radiation is given by (*IMA2* p. 295)

$$P_{\text{rad}} = \frac{1}{3}\alpha T^4 = \frac{4\sigma T^4}{3c}$$

where T is temp in Kelvin, $\alpha = \text{radiation constant} = 4\sigma/c$, $\sigma = \text{Stefan-Boltzmann constant} = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$

Total Pressure

The pressure is the sum of the ideal gas pressure and the radiation pressure (*IMA2* p. 295):

$$P_{\text{total}} = \frac{\rho kT}{\mu m_H} + \frac{1}{3}\alpha T^4$$

A solar model gives a central temperature of $1.57 \times 10^7 \text{ K}$. At this temperature, radiation pressure is $1.53 \times 10^{13} \text{ N m}^{-2}$, only 0.065% of gas pressure and therefore negligible. However, in larger stars radiation pressure plays a major role. A third contributor to pressure, degeneracy pressure, is tiny for the Sun but becomes dominant in degenerate stars.

Gravitation and the Kelvin-Helmholtz Timescale

One likely source of stellar energy is conversion during collapse of gravitational potential energy. A crude calculation assuming a constant density throughout gives an estimated total gravitational energy of

$$U_g \cong -\frac{3}{5} \frac{GM^2}{R}$$

where G is the gravitational constant, M is the total mass of the star, R is the current radius of the star.

The total mechanical energy E of the star by the virial theorem is

$$E \cong -\frac{3}{10} \frac{GM^2}{R}$$

and this same amount of energy must have been radiated away or carried off during collapse (*IMA2* p. 296). The authors show that at constant luminosity L_\odot , the Sun would require only $t_{\text{KH}} = 10^7$ years to emit this much energy, where t_{KH} represents the *Kelvin-Helmholtz Timescale*. This is inconsistent with the estimated age of the solar system based on moon rocks etc., namely approximately 4×10^9 y. (However, gravitational potential energy conversion can play a major role in other situations.)

Nuclear Reactions

I will give only a brief summary of this important and broad subject (see *IMA2* p. 298ff)

For any atom, the number of nuclear protons (*atomic number*) is Z, the number of neutrons N, the total nucleons A (called the *mass number*). The charge on the proton is $+e$, where $-e$ is the charge on the electron.

Atomic particle masses are given by:

$$\begin{aligned}m_p &= 1.67262158 \times 10^{-27} \text{ kg} = 1.0072765 \text{ u} \\m_n &= 1.67492716 \times 10^{-27} \text{ kg} = 1.0086649 \text{ u} \\m_e &= 9.10938188 \times 10^{-31} \text{ kg} = 0.0005485799 \text{ u}\end{aligned}$$

where 1 u has $E = mc^2$ energy equivalence of 931.494013 MeV/c² (using $m = E/c^2$).

The *binding energy* of a nucleus is the positive amount of mechanical energy required to disassemble it into separate parts. A bound system has a lower potential energy than its constituent free parts—this is what keeps the system together. Typically, the difference in mass (*mass deficit*) appears as energy released upon the creation of a bound state (by fusion of nucleons). The reaction releases positive binding energy with the same magnitude as the negative potential energy attained by the fused nucleus.¹²¹ The binding energy of the helium nucleus is 26.731 MeV.

As a rough calculation, conversion of 0.7% of the inner 10% of the Sun's mass of H to He yields

$$E_{\text{nuclear}} = 0.1 \times 0.007 \times Mc^2 = 1.3 \times 10^{44} \text{ J.}$$

This amount of binding energy released at constant luminosity L_\odot for would last $\sim 10^{10}$ years, more than sufficient to power the Sun over the expected age of 4×10^9 y.

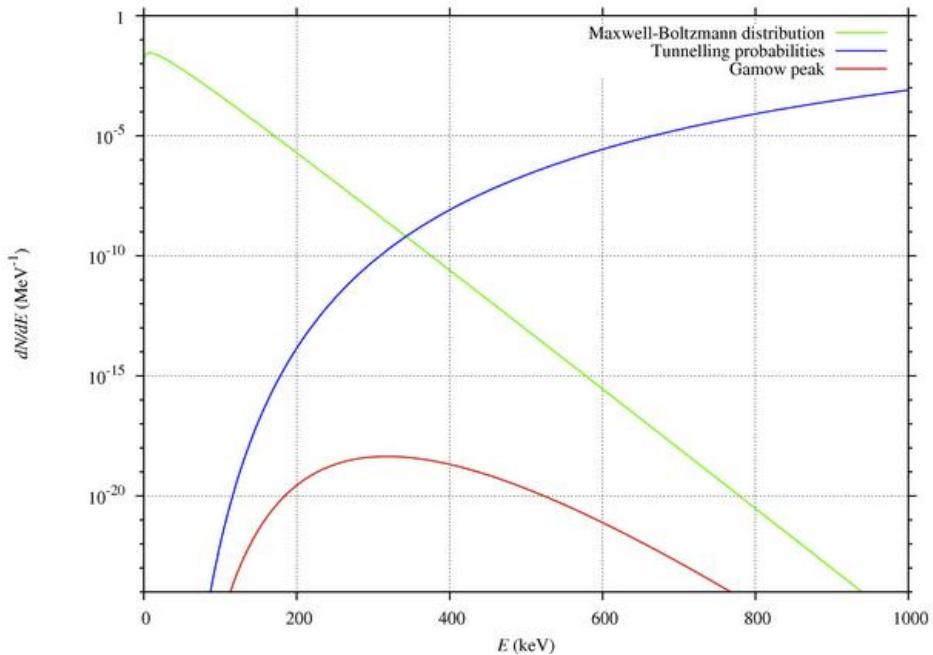
In solar nuclear reactions at expected core temperatures of 1.57×10^7 K, the classical barrier of Coulomb repulsion cannot be overcome to any significant degree. However, *quantum mechanical tunneling* based on *Heisenberg Uncertainty* allows nucleons to overcome this Coulomb repulsion barrier and enter the deep potential well leading to fusion. A rough calculation shows that 1.57×10^7 K is adequate to allow the protons to approach the nucleus by a distance of one proton wavelength, giving the tunneling a realistic probability.

In calculating reaction probabilities, George Gamow (1904 - 1968) showed that a peak probability of reaction exists (the *Gamow peak*) at a particular kinetic energy E of the colliding particles. If E is too low, the nuclear Coulomb barrier cannot be overcome even by tunneling, whereas if E is too high, the $e^{-E/kT}$ Maxwell-Boltzmann distribution term greatly reduces the likelihood that such energetic particles will be present (IMA2 p. 305 and *here*¹²²).

Adjacent Graph: Gamow Peak
(graphed in red) for Fusion
Reaction Probability (fr. Fabrice
Fleurot¹²³)

The Gamow peak occurs at $E_0 = (bkT/2)^{2/3}$, where b is a complicated term formed from the reduced mass, Z_1 and Z_2 , and constants (IMA2 p. 304).

Nuclear reactions also often exhibit *resonances*, typically sharp localized increased reaction probabilities for one or more relatively narrow energy ranges... *Electron screening* by electrons liberated by ionization helps to hide the nucleus and thus reduce the effective Coulomb barrier and improving the reaction rate (by up to 10% to 50%)



¹²¹ http://en.wikipedia.org/wiki/Binding_energy

¹²² Graph and paraphrase from <http://nu.phys.laurentian.ca/~fleurot/fusionrate/>

¹²³ *ibid.*

Nuclear reactions may also be expressed in terms of *power laws* in the form

$$e_{ix} = e'_0 X_i X_x \rho^\alpha T^\beta$$

where e_{ix} is the amount of energy liberated per kg of material, and $X_i X_x$ are the mass fractions of the two particles, and α and β are exponents...

The Luminosity Gradient Equation

To compute the luminosity of a spherically symmetrical star, all of the energy generated by stellar material must be considered:

$$\frac{dL_r}{dr} = 4\pi r^2 \rho \epsilon$$

where L_r is interior luminosity (energy production out to radius r), ϵ is the total energy released per kg per sec by all nuclear reactions and by gravity (the latter may be negative)

Stellar Nucleosynthesis and Conservation Laws

Has Bethe is credited with original discovery in 1938 of details of conversion of H to He in stars as their energy source, although Arthur Eddington (1882–1944) in 1920 had suggested that stars obtained their energy from fusion.

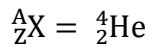
A two-body collision is much more probable than a four-body collision, so that *nucleosynthesis* usually takes place as a series or chain of two-body and one-body reactions.

In each reaction, there must be:

conservation of electric charge (adding positive and negative charges including those of antimatter),
conservation of number of nucleons (neutrons and protons), and
conservation of leptons (i.e., number of leptons minus number of antimatter leptons)
(leptons are electrons, positrons, neutrinos, antineutrinos, muons, taus, not photons)

Neutrinos have a tiny nonzero mass, and have such a low collision probability that almost all created deep inside the Sun escape at the surface (IMA2 p. 309).

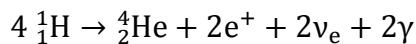
Atoms are represented by the following notation:



where A is total number of nucleons (*atomic mass number*), Z is number of protons (*atomic number*), and X is the element determined by Z.

Proton-Proton Chains

These are one of the main chain of reactions that in sum convert H to He. The PPI branch, which is the most common in the Sun (69%), has reactants and intermediates of ordinary hydrogen or protium (${}_1^1 \text{H}$), deuterium (${}_1^2 \text{H}$) and helium-3 (${}_2^3 \text{He}$) and emissions of positrons, neutrinos, and gamma rays (IMA2 p. 309).



The PPII branch of the PP chain (~31% on the Sun) has reactants and intermediates of helium-3 (${}_2^3 \text{He}$), helium-4 (${}_2^4 \text{He}$), beryllium-7 (${}_4^7 \text{Be}$), and lithium-7 (${}_3^7 \text{Li}$), with emission of electrons, neutrinos, and gamma rays (IMA2 p. 310).

The PPIII branch of the PP chain, which is quite rare on the Sun (<1%), has reactants with electron capture and intermediates of protium ${}_1^1\text{H}$, beryllium-7 (${}_4^7\text{Be}$), boron-8 (${}_5^8\text{B}$), and beryllium-8 (${}_4^8\text{Be}$), with emission of positrons, neutrinos, and gamma rays (IMA2 p. 310).

At the Sun's core temperature (estimated at 1.5×10^7 K), the energy generation rate is proportional to T^4 and the PP chain fusion reactions predominate.

CNO Cycle

This independent cycle, formulated by Hans Bethe (1906 - 2005) in 1938, also converts H to He. The first branch produces carbon-12 and helium-4, has reactants and intermediates of carbon-12 (${}_{12}^6\text{C}$), protium ${}_1^1\text{H}$, nitrogen-13 (${}_{13}^7\text{N}$), carbon-13 (${}_{13}^6\text{C}$), nitrogen-14 (${}_{14}^7\text{N}$), oxygen-15 (${}_{15}^8\text{O}$), nitrogen-15 (${}_{15}^7\text{N}$), and Helium-4 (${}_{2}^4\text{He}$), with emissions of positrons, neutrinos, and gamma rays (IMA2 p. 311).

A second CNO Cycle branch is much less probable. It produces nitrogen-14 and helium-4 and involves reactants and intermediates of carbon-12, protium, nitrogen-13, carbon-13, nitrogen-14, oxygen-15, and nitrogen-15, with emissions of positrons, neutrinos, and gamma rays (IMA2 p. 311).

Comparing PP and CNO cycle reactions, the CNO cycle has a much stronger temperature dependence, with the energy generation rate is proportional to about T^{20} . Low mass stars (less massive than M_\odot) with cooler cores have predominantly PP hydrogen burning, whereas more massive stars (at least somewhat more massive than M_\odot) have H burning dominated by the CNO cycle. CNO cycle burning is seen when the temperature has risen to around 15×10^6 K.

Transition to Helium Burning in the Core

When the stellar core H has been mostly consumed ("burned up"), the mean molecular weight of the gas has increased if T and ρ are unchanged. By the pressure equation of state, if T and ρ are unchanged, the central pressure drops, as given by

$$P_g = \frac{\rho k T}{\mu m_H}$$

This fall in pressure this leads to a partial collapse of the core until the T and density rise to oppose further collapse. When T and density have become sufficiently high, helium nuclei can begin to "burn" (fuse).

Helium burning takes place by the *Triple Alpha Process*, in which three helium-4 nuclei are fused to one carbon-12, with an intermediate of beryllium-8 and emission of gamma rays. This reaction is strongly temperature dependent, with the energy generation rate is proportional to T^{41} . It is seen when the temperature has risen to around $T \sim 100 \times 10^6$ K¹²⁴

Carbon and Oxygen Burning and Nucleosynthesis of Higher Elements

These reactions can also take place in the high temperature environment of helium burning. Carbon-12 and helium-4 combine to form oxygen-16, which combines with helium-4 to produce neon-20, with emission of gamma rays. This reaction takes place with T around 15×10^6 K.

¹²⁴ http://en.wikipedia.org/wiki/Triple-alpha_process

The rising Coulomb barrier makes nucleosynthesis of larger nuclear species less and less likely. But with temperatures rising to the range of 600×10^6 K, it may be possible to synthesize oxygen-16, neon-20, sodium-23, magnesium-23, and magnesium-24.

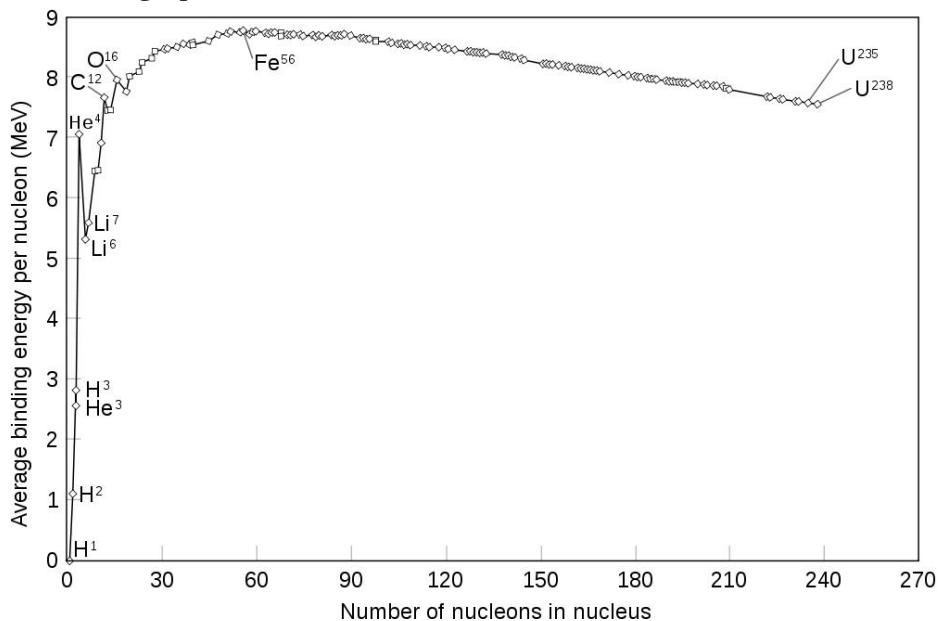
With temperatures around 1000×10^6 K, magnesium-24, silicon-28, phosphorus-31, silicon-31, and sulfur-32 may be formed (IMA2 p. 313).

Some of these reactions are endothermic (require net input of energy), and are more likely to occur in stellar cores than *exothermic* reactions.

Binding Energy Per Nucleon, Nuclear Stability, and Relative Abundances

The binding energy per nucleon is shown in this graph:¹²⁵

Here, the high binding energy per nucleon (BEPN) of Iron-56 is shown, the isotope of greatest nuclear stability. Higher Z atoms have lower BEPN, as do lower-Z elements. Helium-4, carbon-12, and oxygen-16 are unusually stable for their positions and are relatively abundant. Tritium is much less stable than deuterium (and is radioactive). Initially after the Big Bang, most atoms in the universe are thought to have been H or He, but nucleosynthesis within stars and supernovas is thought to have created higher-Z elements throughout the universe.



Energy Transport and Thermodynamics

I will give only highlights of these topics.

Transport of energy from the core to the atmosphere takes place by radiation, convection, and conduction (although conduction is insignificant in most stars during most of their lifetimes).

Radiative Temperature Gradient

The temperature gradient for radiative transport in a spherically symmetrical star is:

$$\frac{dT}{dr} = -\frac{3}{4ac} \frac{\bar{\kappa}\rho}{T^3} \frac{L_r}{4\pi r^2}$$

where the $\bar{\kappa}$ term is the total Rosseland¹²⁶ mean opacity (IMA2 p. 250), a =radiation constant = $4\sigma/c$, ρ is the local density, r is distance from center, L_r is interior luminosity (energy production out to radius r) (IMA2 p. 316).

¹²⁵ http://en.wikipedia.org/wiki/File:Binding_energy_curve_-_common_isotopes.svg

¹²⁶ after Svein Rosseland (1894 - 1985)

Pressure Scale Height H_P

This is a characteristic height of convective regions, and is given by

$$H_P = \frac{P}{\rho g}$$

where P is the pressure, ρ is the local density, and g is the local acceleration of gravity = GM_r/r^2 . A typical value for the Sun is $\sim R_\odot/10$.

Formulas for *First Law of Thermodynamics* (IMA2 p. 318), *Specific Heats* C_V and $C_P = C_V + nR$, parameter $\gamma = C_P/C_V$, and the adiabatic gas law are omitted here.

Adiabatic Speed of Sound

This derives from the compressibility of a gas and its inertia (represented by density):

$$v_s = \sqrt{\gamma P/\rho}$$

Adiabatic Temperature Gradient

This describes how the temperature in a gas bubble changes as it rises from convection and expands adiabatically. It is given by:

$$\left. \frac{dT}{dr} \right|_{\text{adiabatic}} = -(1 - \frac{1}{\gamma}) \frac{\mu m_H}{k} \frac{GM_4}{4r^2}$$

where k = Boltzmann constant, $\gamma = C_P/C_V$, etc. [see earlier this chapter]. If the actual temperature gradient is steeper than this adiabatic gradient (a *superadiabatic gradient*), nearly all of the interior luminosity may be carried in this region by convection. Thus, either radiation or convection often predominate in a particular region, depending on the temperature gradient (IMA2 p. 322).

Criteria for Convection

The criteria derived for convection of a hot gas bubble are:

$$\left| \frac{dT}{dr} \right|_{\text{actual}} > \left| \frac{dT}{dr} \right|_{\text{adiabatic}}$$

or equivalently,

$$\frac{d \ln P}{d \ln T} < \frac{\gamma}{\gamma - 1}$$

where again $\gamma = C_P/C_V$. For a monoatomic gas, $\gamma=5/3$ and convection occurs where the LHS of the 2nd equation is < 2.5 . If the LHS is > 2.5 , the region is stable against convection.

Convection tends to occur when one or more of these apply:

- (1) the stellar opacity is large (limiting radiative transport because an unachievable steep temp. gradient would be required),
- (2) a region exists where ionization is occurring, causing a large specific heat and a low adiabatic temperature gradient, and/or
- (3) the temperature dependence of the nuclear energy generation rate is large, causing a steep radiative flux gradient and a large temperature gradient.

Items (1) and (2) can occur in the atmospheres of stars simultaneously, whereas (3) only occurs deep in stellar interiors when CNO or triple-alpha processes are occurring (IMA2 p. 325).

Mixing-Length Theory of Superadiabatic Convection

The authors show that the temp gradient need be only slightly superadiabatic in the deep interior in order for convection to carry most of the energy. The derivation introduces the *mixing length* l , which represents the distance a hot bubble must rise before it *thermalizes* (its excess heat/temperature is dissipated to the surrounding gas). H_P is the pressure scale height (distance for pressure to drop to $1/e = 0.37$ of the highest or inner value)... The *convective flux* F_C is derived. Mixing-Length Theory uses arbitrary empirical constants but has been relatively successful in predicting the results of observations.

In an example calculation, a *convective velocity* for the Sun at $r=0.714$ is estimated at 50 m s^{-1} , which is 10^{-4} vs.(the local solar sound speed).

Stellar Model Building

Section discusses Entropy, Constitutive Relations, Boundary Conditions, the Vogt-Russell Theorem, Numerical Modeling, Polytropic Models (special cases which allow analytical solutions), and Poisson's Equation. The highly idealized Lane-Emden Equation is derived (*IMA2* p. 336).

The Main Sequence

As a result of H burning in the stellar core, the Vogt-Russell Theorem requires that such a change of the composition or mass of a star leads to adjustments in the effective temperature and luminosity. In other words, what can be observed arising from the star's surface reflects the state of the star deep within. As stellar mass increases, the central P and T increase.

Lower Mass Stars and Lower Main Sequence Mass Limit

As initial star mass decreases, a lower limit is reached at which point the central hydrogen-burning nuclear reactions (if ever present) are insufficient to maintain a sufficiently high temperature to stabilize the star against gravitational contraction (see later discussions). Such contracted stars are no longer on the main sequence. This lower mass limit is about $0.072 M_\odot$ (for Population I having high Z)—it is somewhat higher, 0.09 , for Population II stars with low metallicity $Z \approx 0$. (*IMA2* p. 341)

For lower mass main sequence stars ($< 1.2 M_\odot$), radiation is adequate in the core to transport the energy released and there is no core convection. However, convection forms near the surface due to the fall of effective temperature with decreasing mass, causing increasing opacity arises partly from H ionization, which makes convection more efficient. The inner limit of convection descends closer to the center with decreasing mass, until at $0.2 M_\odot$ and smaller, the entire star is convective. (*IMA2* p. 343)

Higher Mass Stars and Upper Main Sequence Mass Limit (Eddington Luminosity Limit)

Larger stars have higher luminosities and at an upper limit in size and luminosity, the radiation pressure can dominate over the gas pressure in certain regions, often in the outer layers. Arthur Eddington (1882 - 1944) gave us the Eddington limit on luminosity:

$$L_{Ed} = \frac{4\pi Gc}{\bar{\kappa}} M$$

where M is the star's mass, and the kappa term is the total Rosseland mean opacity (for which larger values signify greater opacity and more rapid intensity attenuation). This is the maximum radiative luminosity that a star can have and still remain in hydrostatic equilibrium. Beyond this luminosity, the envelopes of massive stars become more loosely bound, and such larger stars suffer large mass loss from being “blown off” as

stellar wind (*IMA2* p. 343 and *here*¹²⁷). The Eddington limit for pure H II is about 3.2×10^4 (M/M_\odot) L_\odot . Eddington's formula has been updated, does not require spherical symmetry, and is now applied to novae, gamma ray burst, and accretion disks as well as to stars. Stars of about $90 M_\odot$ are near this limit. High mass main sequence stars ($> 90 M_\odot$) exhibit rapid mass loss and often oscillatory or variable nuclear energy generation.

The term Eddington accretion rate (an analogous limit) is applied to accretion disks, see Chapter 18.

For larger mass main sequence stars, convection is dominant at their cores because the CNO-cycle burning is strongly temperature dependent producing a strong temp gradient, plus radiation is not efficient enough to transport the rate of energy release. Outside of the hydrogen burning core, convection ceases as radiation can handle the transport. (*IMA2* p. 343) Stars less than $10 M_\odot$ may also have an outer convection zone.¹²⁸

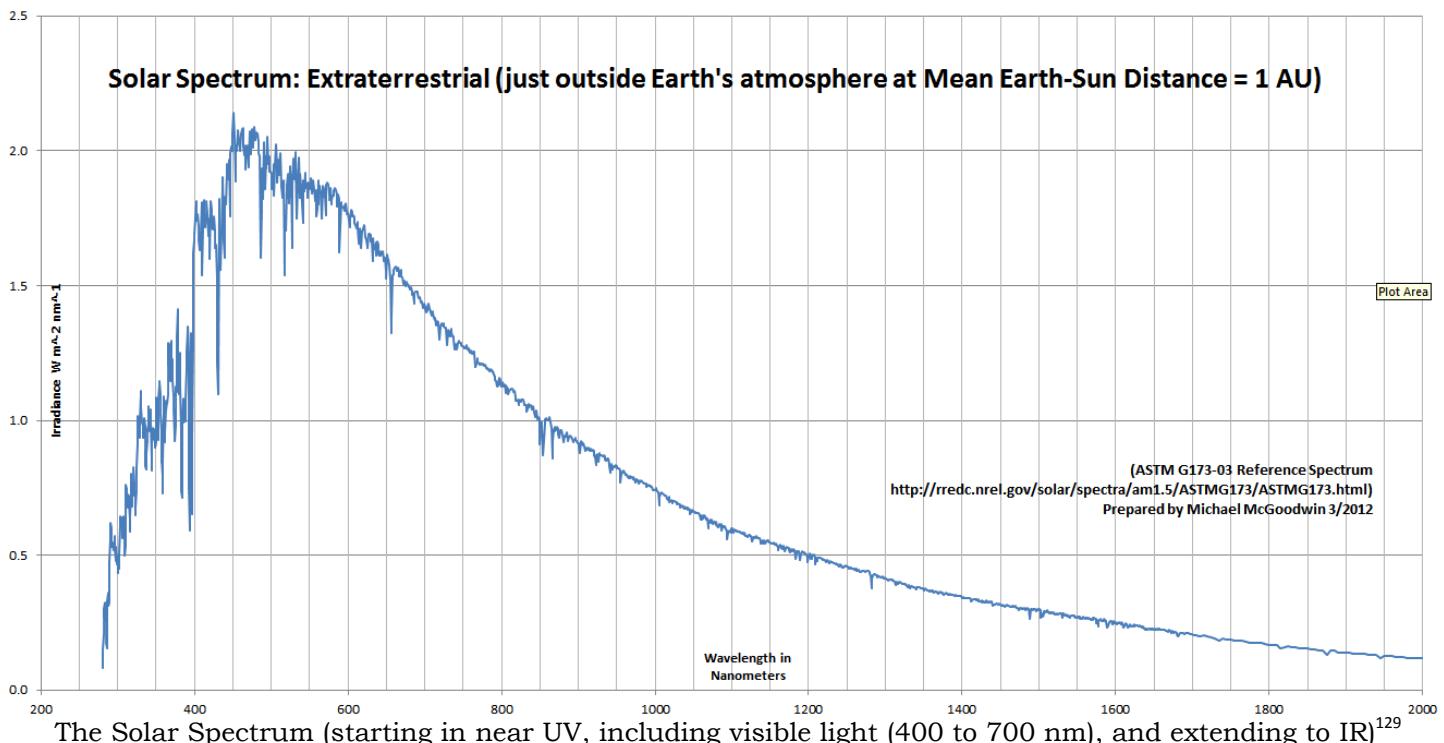
Range of Luminosities and Lifetimes

Over the range of masses of main sequence stars, luminosities range from a lower limit of about $5 \times 10^{-4} L_\odot$ ($T_{\text{eff}} \approx 1700\text{K}$) to an upper limit of about $10^6 L_\odot$ ($T_{\text{eff}} \approx 53,000\text{K}$), spanning nine orders of magnitude. The much greater luminosities of massive stars are associated with correspondingly much shorter lifetimes. Of course, the substantial differences in effective temperatures are associated with characteristic changes in the stellar spectra.

¹²⁷ http://en.wikipedia.org/wiki/Eddington_luminosity

¹²⁸ http://en.wikipedia.org/wiki/Convection_zone

The Sun (Chapter 11)



Note: I have also presented substantial information about the Sun in a Space Physics summary *here*.¹³⁰

The Solar Interior

Our star, the Sun, is divided into an *atmosphere* (*photosphere, chromosphere, transition region, and corona*) and an *interior* (*radiative and convection zones*). Going outward, the base of the atmosphere may be taken to be the region in which optical depth has fallen to less than unity. (See below for a more exact definition of the base of the photosphere.) The optical depth at wavelength λ for a path through a stellar atmosphere is given by

$$\frac{I}{I_0} = e^{-\tau_\lambda}$$

For an optical depth ≤ 1 , the intensity of light at wavelength λ decreases to $\geq e^{-1} = 1/e = 0.368$ of the initial intensity value. At an optical depth = 2, a reduction to $1/e^2 = 0.135$ of the initial intensity value occurs.

Evolutionary History of the Sun

The Sun is spectral class G2 luminosity class V star. The surface (atmospheric) composition is X = 0.74, Y = 0.24, and Z = 0.02 (where XYZ are the mass fractions of H, He, and “metals”, resp.) Comparison with the

¹²⁹ <http://rredc.nrel.gov/solar/spectra/am1.5/ASTMG173/ASTMG173.html>.

Reference solar spectral irradiance data and graphs are available here, including the graphable Excel data shown above. Graph created by MCM.

¹³⁰ http://www.mcgoodwin.net/pages/spacephysics_ess471.pdf

unchanging moon and meteoritic rocks places the Sun's age is estimated at 4.57 Ga (Ga = Gy = 10^9 Julian years, each of which is 31557600 SI seconds in length).¹³¹

According to stellar models,¹³² during the Sun's evolution as a main sequence star over 4.57 Ga,

- the luminosity L of the Sun has steadily risen from an initial value of $0.677 L_\odot$ to $1 L_\odot$ at the present. It will continue to rise at least to 8 Ga.
- the radius of the Sun R has steadily risen from an initial value of $0.869 R_\odot$ to $1 R_\odot$ currently. It will continue to rise at least to 8 Ga.
- the effective temperature of the Sun t_{eff} has steadily risen from an initial value of 5620 K to 5777K currently. The model shows that it will rise further and attain a peak near 5800 K at around 6.5 Ga, then drop to 5780 K by 8 Ga.

The impact of these changes on Earth depends on complex factors including Earth atmospheric convection, Earth atmospheric composition, and Earth atmospheric and ground albedo.

The Present Interior of the Sun

The textbook presents considerably more detail than summarized here, including useful graphs of distributions of P, T, density, etc.

According to the 2001 models of Bahcall, Pinsonneault, & Basu, the core temperature is 1.57×10^7 K and this drops steadily outward. The core pressure is 2.342×10^{16} N m⁻², and this falls off more rapidly with distance than does T. The core density 1.527×10^5 kg m⁻³ and this also falls off rapidly similar to P. The current core mass fractions are X = 0.3397 (H has decreased from initial 0.71) and Y = 0.6405 (He has increased from 0.27). Near the surface, gravitational settling has decreased Y by 0.03 and enriched X by 0.03.

The base of the present day solar convection zone (the *tachocline*¹³³) is at $0.714 R_\odot$. Below this level, the Sun is purely radiative. Convection arises when the temperature gradient is superadiabatic,

$$\left| \frac{dT}{dr} \right|_{\text{actual}} > \left| \frac{dT}{dr} \right|_{\text{adiabatic}}$$

and for reasons previously discussed (see Chapter 10). (For more massive stars, convection begins in the center of the star, Chapter 13).

The peak rate of energy production (luminosity) from fusion is seen at about $0.1 R_\odot$. (This result partly reflects the amount of mass contained in a tiny shell at the center vs. a much greater mass in a shell of identical width at $0.1 R_\odot$, and also reflects the maximal depletion of H in the center.) Luminosity does not increase beyond about $0.3 R_\odot$ because there is the temperature is too low for effective fusion.

90% of the mass of the sun is enclosed at about $0.5 R_\odot$.

Differential Rotation of the Interior

The rotation rate of the interior of the Sun varies with depth. Deep to the tachocline (in the radiation zone), the rotation angular velocity is essentially constant with changing r/R_\odot . However, in the convection zone of the interior lying above the tachocline, the rate of rotation varies with both latitude and depth.¹³⁴ The shear at

¹³¹ Bouvier A, Meenakshi W, "The age of the solar system redefined by the oldest Pb[207]-Pb[206] age of a meteoritic inclusion". *Nature Geoscience*, Published online: 22 August 2010

<http://www.nature.com/ngeo/journal/v3/n9/full/ngeo941.html>

¹³² Bahcall HN, Pinsonneault MH, & Basu S, "Solar Models: Current Epoch and Time Dependences, Neutrinos, and Helioseismological Properties", *The Astrophysical Journal*, 555:990-1012, 2001 July 10, <http://iopscience.iop.org/0004-637X/555/2/990/fulltext/52957.text.html> and IMA2 p. 350

¹³³ Tachocline:

- <http://solarscience.msfc.nasa.gov/interior.shtml>
- <http://en.wikipedia.org/wiki/Tachocline>

¹³⁴ http://en.wikipedia.org/wiki/Differential_rotation

the tachocline is thought to generate electric currents and play a role in generation of the magnetic field (IMA2 p. 364).

The Solar Neutrino Problem

Experiments by Raymond Davis and John Bahcall (using data gathered deep in the Homestake Gold Mine), and subsequent experiments by Japan's Super-Kamiokande observatory¹³⁵ and SAGE and GALLEX experiments, measured rates of neutrinos generated in the Sun's core and captured in the experimental system. Davis expressed reaction rates as solar neutrino units (SNUs), defined as 10^{-36} reactions per second per target atom (Chlorine-37 conversion to radioactive Argon-37, in Davis's case). They consistently found insufficient SNUs according to the then prevailing models. The explanation came with *neutrino oscillations* in the Mikheyev-Smirnov-Wolfenstein effect, changing our understanding of particle physics.¹³⁶

The Solar Atmosphere

The Sun's apparently *sharp edge* (in visible light) is actually formed in a region 600 km thick (0.09% of R_\odot) over which the atmosphere changes rapidly from *optically thick* (small optical depth) to *optically thin* (larger optical depth).

Photosphere

Because of the wavelength-dependence of optical depth, the base of the photosphere is defined relative to the level where the optical depth $\tau_\lambda = \tau_{500 \text{ nm}} = 1$. That is, the τ_λ of the atmosphere at a wavelength of 500 nm is 1, and the intensity of a light ray is reduced to $1/e = 0.368$. The *base of the photosphere* is arbitrarily taken to be 100 km *deeper* than the $\tau_{500 \text{ nm}} = 1$ level. At this only slightly deeper level, the optical depth is much greater ($\tau_{500 \text{ nm}} = 23.6$), almost no light can penetrate by radiation, and the temperature is about 9400 K.

The photosphere is where most of the visible (optical) photons we observe originate (very few come from below the base as defined). Temperature in the photosphere decreases with increasing radius from the Sun center, until one reaches a minimum value of about 4400 K at the *outer limit of the photosphere*. This region of minimum temperature is about 525 km above the $\tau_{500 \text{ nm}} = 1$ level or 525 km above the photosphere base as defined above. Thus, the entire photosphere is about 625 km thick.

The *effective temperature* of the photosphere, which we measure to be about 5777 K, derives from the average level from which photons that reach us have arisen. In the Eddington approximation, this proves to be where $\tau = 2/3$ and this corresponds to $T_{\text{eff}} = T_{\tau = 2/3} = 5777$ K, part way between the extremes of 9400 K and 4400 K.

The visible and IR solar spectrum resembles that of a blackbody continuum, consistent with a source of opacity for these bands. The opacity is thought to arise from the presence of H⁻ ions (a hydrogen atom with extra electron)—although they are rare, they are much more likely than H I to contribute to continuum radiation. Different parts of the solar spectrum may arise from different levels in the photosphere.

Absorption lines are also produced in the photosphere. The darkest central part of the lines are produced in the cooler outer regions of the photosphere. Off center from the dark peak, the lesser absorption arises from deeper hotter levels (MCM: presumably because of higher velocities and greater Doppler spread of wavelengths).

Granules: I have discussed *solar granulations* and other Sun topics [here](#).¹³⁷ Granulations arise from convecting cells spanning about 700 km and lasting 5 to 10 minutes. Spectra obtained across these granulations show wiggles of the lines from Doppler shifts due to the rising (blueshifted, brighter, and more

¹³⁵ <http://www-sk.icrr.u-tokyo.ac.jp/sk/index-e.html>

¹³⁶ Neutrino oscillations:

- http://www.nobelprize.org/nobel_prizes/physics/articles/bahcall/
- http://en.wikipedia.org/wiki/Solar_neutrino_problem

¹³⁷ http://www.mcgoodwin.net/pages/spacephysics_ess471.pdf

central) and descending (redshifted, darker, and more peripheral) portions of gas. “Granules are small (about 1000 km across) cellular features that cover the entire Sun except for those areas covered by sunspots. These features are the tops of convection cells where hot fluid [gas] rises up from the interior in the bright areas, spreads out across the surface, cools and then sinks inward along the dark lanes. Individual granules last for only about 20 minutes. The granulation pattern is continually evolving as old granules are pushed aside by newly emerging ones... The flow within the granules can reach supersonic speeds of more than 7 km/s ... and produce sonic booms and other noise that generates waves on the Sun's surface.”¹³⁸

Faculae: “*Faculae* are bright areas [in the photosphere] that are usually most easily seen near the limb, or edge, of the solar disk. These are also magnetic areas but the magnetic field is concentrated in much smaller bundles than in sunspots. While the sunspots tend to make the Sun look darker, the faculae make it look brighter. During a sunspot cycle the faculae actually win out over the sunspots and make the Sun appear slightly (about 0.1%) brighter at sunspot maximum than at sunspot minimum.”¹³⁹

Differential Rotation of the Photosphere

The rotation rate (angular velocity) of the gas of the photosphere also varies with *latitude*, from a period of about 25 days near the equator to about 36 days near the poles. I have discussed *solar differential rotation* in relation to magnetic fields and sunspots [here](#).¹⁴⁰

Chromosphere

This “sphere of color” lies just above the photosphere, starting therefore at about 525 km above the $\tau_{500 \text{ nm}} = 1$ level and extending about 1600 km to 2100 km above the $\tau_{500 \text{ nm}} = 1$ level. The gas density drops by a factor of 10^4 over this distance, and the temperature gradually rises from 4400 K at the bottom to about 10,000 K at the top. This region imparts lines to the solar spectrum that are not produced in the photosphere, particularly H Balmer lines and those of He II, Fe II, Si II, Cr II, and Ca II (especially the H 396.8 nm line and K 393.3 lines). The lines added include some *emission lines* arising in the outer hot zone, according to Kirchhoff’s second law of spectroscopy.

The chromosphere is not usually visible, as it is overwhelmed by the photosphere, but can be seen in a solar eclipse as a reddish region. In addition, the *flash spectrum* as the Sun’s upper limb just rises or sets appears as red from the Balmer Ha 656.3 nm emission line arising in the chromosphere.¹⁴¹ Imaging using the Balmer Ha emission line reveals *supergranulation* on a scale of 30,000 km in the solar atmosphere, apparently reflecting underlying convective phenomena (*IMA2* p. 365 and [here](#)).¹⁴²

The chromosphere may also be imaged using the 393.4 nm emission line of Ca II [Ca II K line]. When the Sun is viewed through a spectrograph or a filter that isolates the H-alpha emission, a wealth of new features can be seen. These features include the *chromospheric network* of magnetic field elements, bright plage around sunspots, dark filaments across the disk and prominences above the limb.¹⁴³

Spicules: *Spicules* are vertically-oriented filaments of rapidly ascending gas (15 km s^{-1}) extending as much as 10,000 km or more above the usual level of the outer chromosphere into the corona.¹⁴⁴ These are relatively small but widespread eruptions, and thus they eject substantial material into the corona. These are also best observed in the Ha emission line. “They last but a few minutes but in the process eject material off of the surface and outward into the hot corona at speeds of 20 to 30 km/s.”¹⁴⁵

¹³⁸ <http://solarscience.msfc.nasa.gov/feature1.shtml>

¹³⁹ ibid

¹⁴⁰ ibid.

¹⁴¹ <http://www.phys.vt.edu/~heremans/Astrolab1156/Readings/flash.html>

¹⁴² Solar Supergranulations:

- <http://solarphysics.livingreviews.org/Articles/lrsp-2010-2/download/lrsp-2010-2Color.pdf>
- <http://solarscience.msfc.nasa.gov/feature1.shtml> (this author lists supergranulation under photospheric phenomena rather than chromospheric)

¹⁴³ <http://solarscience.msfc.nasa.gov/chromos.shtml>

¹⁴⁴ <http://history.nasa.gov/SP-402/p2.htm>

¹⁴⁵ <http://solarscience.msfc.nasa.gov/feature2.shtml>

Plage: “A *plage* (from Fr. ‘beach’) is a bright region in the chromosphere of the Sun, typically found in regions of the chromosphere near sunspots. The plage regions map closely to the faculae in the photosphere below...”

¹⁴⁶ “*Plage* ... are bright patches [in the chromosphere] surrounding sunspots that are best seen in H-alpha.

Plage are also associated with concentrations of magnetic fields and form a part of the network of bright emissions that characterize the chromosphere.”¹⁴⁷

Filaments: “*Filaments* are dark, thread-like [chromospheric] features seen in the red light of hydrogen (H-alpha). These are dense, somewhat cooler, clouds of material that are suspended above the solar surface by loops of magnetic field.”¹⁴⁸

Prominences: “Prominences are dense clouds of material suspended above the surface of the Sun by loops of magnetic field. Prominences and filaments are actually the same things except that prominences are seen projecting out above the limb, or edge, of the Sun. Both filaments and prominences can remain in a quiet or quiescent state for days or weeks. However, as the magnetic loops that support them slowly change, filaments and prominences can erupt and rise off of the Sun over the course of a few minutes or hours.”¹⁴⁹

“A prominence is a large, bright feature extending outward from the Sun's surface, often in a loop shape. Prominences are anchored to the Sun's surface in the photosphere, and extend outwards into the Sun's corona. While the corona consists of extremely hot ionized gases, known as plasma, which do not emit much visible light, prominences contain much cooler plasma, similar in composition to that of the chromosphere... A typical prominence extends over many thousands of kilometers; the largest on record was observed by the Solar and Heliospheric Observatory (SOHO) in 2010 and is estimated at over 700,000 kilometers... long – roughly the radius of the Sun... A prominence forms over timescales of about a day, and stable prominences may persist in the corona for *several months*. Some prominences break apart and give rise to coronal mass ejections.”¹⁵⁰

The Transition Region

The temperature rises rapidly at the top of the chromosphere in the *transition region*, attaining 10^5 K in only 100 km and about 10^6 K in 200 km. It can be seen in the Ly-a emission line, the CIII 97.7 nm line, and other UV emission lines seen at high temperatures (IMA2 p. 365). “The transition region does not occur at a well-defined altitude. Rather, it forms a kind of nimbus around chromospheric features such as spicules and filaments [that extend into the corona], and is in constant, chaotic motion. The transition region is not easily visible from Earth's surface, but is readily observable from space by instruments sensitive to the extreme ultraviolet portion of the spectrum.”¹⁵¹

“The transition region is a thin and very irregular layer of the Sun's atmosphere that separates the hot corona from the much cooler chromosphere. Heat flows down from the corona into the chromosphere and in the process produces this thin region where the temperature changes rapidly from 1,000,000°C ... down to about 20,000°C... Hydrogen is ionized ... at these temperatures and is therefore difficult to see. Instead of hydrogen, the light emitted by the transition region is dominated by such ions as C IV, O IV, and Si IV... These ions emit light in the ultraviolet region of the solar spectrum that is only accessible from space.”¹⁵²

The Corona

The corona (“crown”), a plasma atmosphere located above the transition region, becomes visible only in solar eclipses. Its energy output is 10^{-6} that of the photosphere. It is extremely tenuous, having a particle number density of 10^{15} m^{-3} at the base, thinning to 10^{15} m^{-3} in the solar wind arriving at the Earth. It is not in local thermodynamic equilibrium, so temperatures are not strictly definable. However, the presence of highly

¹⁴⁶ http://en.wikipedia.org/wiki/Plage_%28astronomy%29

¹⁴⁷ <http://solarscience.msfc.nasa.gov/feature2.shtml>

¹⁴⁸ ibid

¹⁴⁹ ibid

¹⁵⁰ http://en.wikipedia.org/wiki/Solar_prominence

¹⁵¹ <http://en.wikipedia.org/wiki/Sun#Atmosphere>

¹⁵² http://solarscience.msfc.nasa.gov/t_region.shtml

ionized iron (Fe-XIV) lines accompanied by substantial Doppler line broadening indicates a plasma temperature in excess of 1 to 2 $\times 10^6$ K.

"The corona is the extended outer atmosphere of the Sun, which is much larger in volume than the Sun itself. The corona continuously expands into space forming the solar wind, which fills all the Solar System. The low corona, near the surface of the Sun, has a particle density around 10^{15} – 10^{16} m⁻³. The average temperature of the corona and solar wind is about 1,000,000–2,000,000 K; however, in the hottest regions it is 8,000,000–20,000,000 K. While no complete theory yet exists to account for the temperature of the corona, at least some of its heat is known to be from magnetic reconnection."¹⁵³

Coronal Light Components: Coronal radiation is subdivided into three partly coextensive components:

- the *K corona*—Kontinuierlich=continual white light arising between 1 and 2.3 R_☉ from scattering of sunlight off free electrons,
- the *F corona*—named after Fraunhofer, but describing scattering from dust grains beyond 2.3 R_☉, a phenomenon which merges with the *zodiacal light* (emitted from the *zodiacal cloud* of dust), and
- the *E corona*—Emission lines from highly ionized atoms, including *forbidden transitions* from metastable states observable due to low number densities and the rarity of collisions.

Radio Emission: The corona is a source of *radio-wavelength non-blackbody radiation*, which arises from free-free electron transitions (aka *bremsstrahlung* or *braking radiation* arising from electron-ion near-collisions). Wavelengths of 1 to 20 cm are observed from the outer corona (mixed in with synchrotron radiation also arising in the corona). (*IMA2* p. 369 and *here*¹⁵⁴)

X-ray Emission: The corona is also a source of X-rays, again not arising as blackbody radiation but rather from multiply-ionized elements such as iron, oxygen, and neon. Wavelengths illustrated in *IMA2* are 14 - 17 nm. X-ray emission is not uniform across the Sun, but arises in active hot regions. Weaker emissions arise from darker and cooler *coronal holes*... (*IMA2* p. 370) "The corona shines brightly in x-rays because of its high temperature. On the other hand, the "cool" solar photosphere emits very few x-rays. This allows us to view the corona across the disk of the Sun when we observe the Sun in X-rays... In the early 70's Skylab carried an x-ray telescope that revealed coronal holes and coronal bright points for the first time. During the last decade, Yohkoh provided a wealth of information and images on the sun's corona. Today we have the SOHO and TRACE [Transition Region and Coronal Explorer] satellites obtaining new and exciting observations of the Sun's corona, its features, and its dynamic character."¹⁵⁵

Helmet streamers: "Helmet streamers are large cap-like coronal structures with long pointed peaks that usually overlie sunspots and active regions. We often find a prominence or filament lying at the base of these structures. Helmet streamers are formed by a network of magnetic loops that connect the sunspots in active regions and help suspend the prominence material above the solar surface. The closed magnetic field lines trap the electrically charged coronal gases to form these relatively dense structures. The pointed peaks are formed by the action of the solar wind blowing away from the Sun in the spaces between the streamers."¹⁵⁶

Coronal loops: "Coronal loops are found around sunspots and in active regions. These structures are associated with the closed magnetic field lines that connect magnetic regions on the solar surface. Many coronal loops last for days or weeks but most change quite rapidly... Some loops, however, are associated with solar flares and are visible for much shorter periods. These loops contain denser material than their surroundings. The three-dimensional structure and the dynamics of these loops is an area of active research."¹⁵⁷

Coronal Holes: "Coronal holes are regions where the corona is dark. These features were discovered when X-ray telescopes were first flown above the earth's atmosphere to reveal the structure of the corona across the solar disc. Coronal holes are associated with "open" magnetic field lines and are often found at the Sun's poles. The high-speed solar wind is known to originate in coronal holes."¹⁵⁸

¹⁵³ <http://en.wikipedia.org/wiki/Sun#Atmosphere>

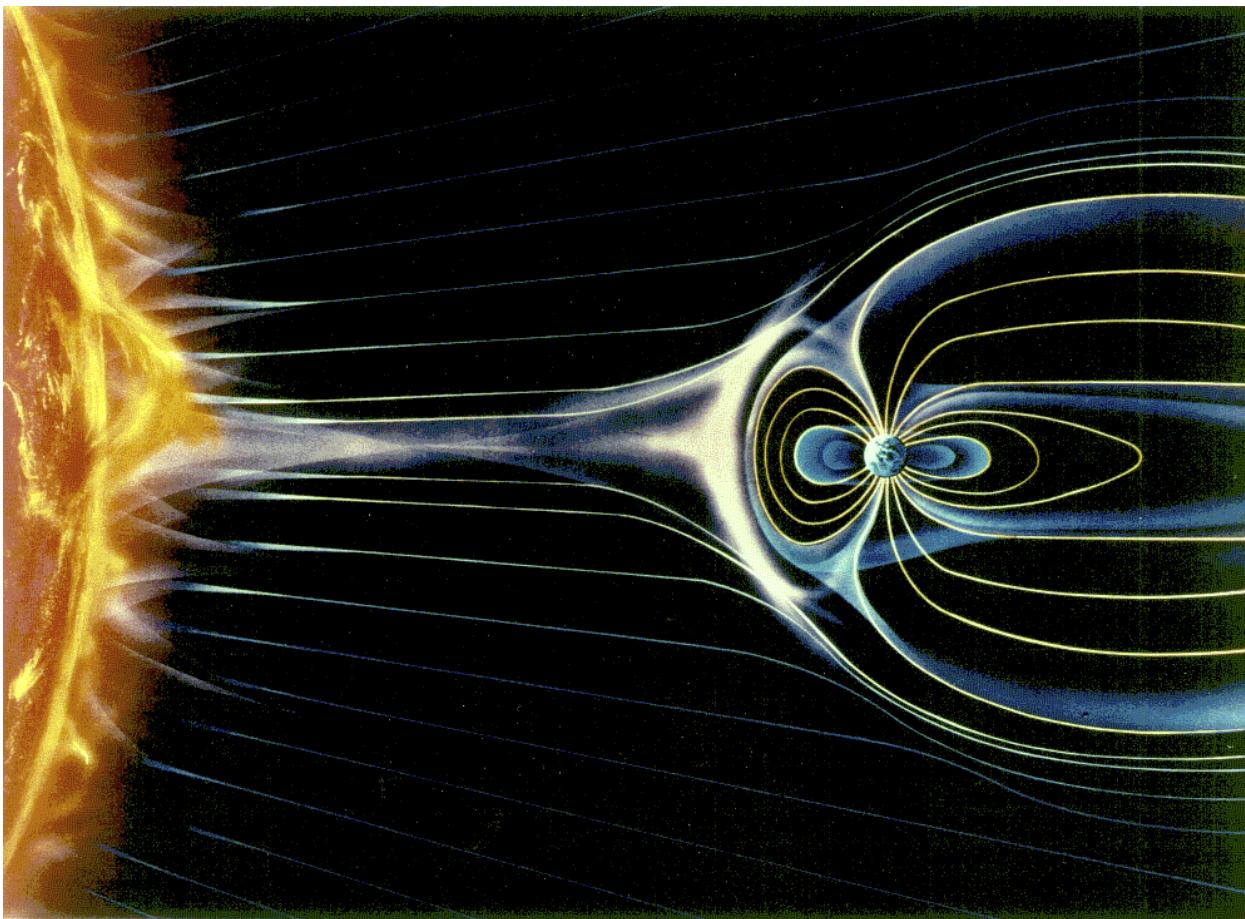
¹⁵⁴ <http://www.cv.nrao.edu/course/astr534/FreeFreeEmission.html>

¹⁵⁵ <http://solarscience.msfc.nasa.gov/corona.shtml>

¹⁵⁶ <http://solarscience.msfc.nasa.gov/feature3.shtml>

¹⁵⁷ ibid.

¹⁵⁸ ibid.



Model of chromospheric and coronal spicules,
and solar wind interacting with Earth magnetosphere (image from NASA)¹⁵⁹

Solar Flares

“A solar flare is a sudden brightening observed over the Sun surface or the solar limb, which is interpreted as a large energy release of up to 6×10^{25} joules of energy (about a sixth of the total energy output of the Sun each second)... The flare ejects clouds of electrons, ions, and atoms through the corona into space. These clouds typically reach Earth a day or two after the event. The term is also used to refer to similar phenomena in other stars, where the term stellar flare applies... The same energy releases may produce coronal mass ejections (CME), although the relation between CMEs and flares is still not well established.”¹⁶⁰

“In a matter of just a few minutes [solar flares] heat material to many millions of degrees and release as much energy as a billion megatons of TNT. They occur near sunspots, usually along the dividing line (neutral line) between areas of oppositely directed magnetic fields... Flares release energy in many forms - ... Gamma rays and X-rays, energetic particles (protons and electrons), and mass flows. Flares are characterized by their ... X-ray flux. The biggest flares are *X-Class flares* [and have peak flux at 100 - 800 pm of $>10^{-4} \text{ W m}^{-2}$]. M-Class flares have [$10^{-5} - 10^{-4} \text{ W m}^{-2}$], and C-Class flares have [$10^{-6} - 10^{-5} \text{ W m}^{-2}$]... The key to understanding and predicting solar flares is the structure of the magnetic field around sunspots. If this structure becomes twisted and sheared then magnetic field lines can cross and reconnect with the explosive release of energy... The flare (the bright area [see image]) lies along a section of a neutral line where the magnetic field is twisted

¹⁵⁹ http://solarscience.msfc.nasa.gov/images/sunearth_lg.gif

¹⁶⁰ http://en.wikipedia.org/wiki/Solar_flare

(or sheared) to point along the neutral line instead of across it. We have found that this shear is a key ingredient in the production of solar flares.”¹⁶¹

Coronal Mass Ejections (CMEs)

“A coronal mass ejection (CME) is a massive burst of solar wind, other light isotope plasma, and magnetic fields rising above the solar corona or being released into space... Coronal mass ejections are often associated with other forms of solar activity, most notably solar flares, but a causal relationship has not been established. Most ejections originate from active regions on Sun’s surface, such as groupings of sunspots associated with frequent flares. Near solar maxima the Sun produces about three CMEs every day, whereas near solar minima there is about one CME every five days.”¹⁶²

Comparison of Major Solar Outbursts

Phenomenon	Onset & Duration	Energy &/or Size	Location	Comments
Solar Flare	Explosively sudden—lasting milliseconds to > 1 hr	Up to 6×10^{25} J as UV, X-ray, gamma ray & all other EM rad.; \pm protons; up to 100,000 km	affect all layers	Categorize by X-ray flux; may arise from magnetic reconnection or sigmoids ¹⁶³ of energy
Solar Prominence	1 day onset; lasts days to months; stable and/or eruptive	up to $1 R_{\odot}$	arise in photosphere, affect all layers	~ Loop shape, some give rise to CMEs
Coronal Mass Ejection ¹⁶⁴	onset over hours	released matter averages 5×10^{13} kg at 20 to 3200 km/s as hot plasma + EM radiation	usu. arise in active sunspot regions, some associated with eruptive prominences; cross all layers	some but not all assoc. w large explosive solar flares; cause auroras; may arise from magnetic reconnection

Miscellaneous Solar Topics

Omitted Solar Topics

I have omitted summarization of the following topics:

Alfvén waves (transverse MHD waves) and their speed ($\approx 10 \text{ m s}^{-1}$)¹⁶⁵

Auroras

Heliopause and heliosheath

Helioseismology

¹⁶¹ Solar Flares:

- <http://solarscience.msfc.nasa.gov/flares.shtml>
- http://en.wikipedia.org/wiki/Solar_flare (quantitation of flux)

¹⁶² http://en.wikipedia.org/wiki/Coronal_mass_ejections

¹⁶³ <http://www.astromynow.com/090420energyribbons.html>

¹⁶⁴ <http://www-istp.gsfc.nasa.gov/istp/nicky/cme-chase.html>

¹⁶⁵ http://en.wikipedia.org/wiki/Alfv%C3%A9n_wave

Lorentz force ($\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$) (which follows the RH rule where v =index, B =middle finger, F =thumb)
 Magnetic dynamo model of the solar cycle¹⁶⁶
 Magnetic energy density
 Magnetic pressure gradient
 Magnetic reconnections¹⁶⁷
 Magnetohydrodynamics MHD
 Parker Solar Wind model (Heliospheric Current Sheet, Parker Spiral, Garden Hose Angle, and Ballerina Skirt)
 Solar cosmic rays
 Solar magnetic fields
 Solar wind and its Fast and Slow components¹⁶⁸

as these have been discussed in another course summary I wrote¹⁶⁹ or will be the subjects of more intensive future courses I hope to take.

Solar Rotation Slowing

Stars like the Sun are observed to be slowing down in rotation rate. Late main-sequence stars with convective envelopes have much slower rotation rates than stars on the upper end of the main sequence. This may be due to transfer of angular momentum to solar wind.¹⁷⁰

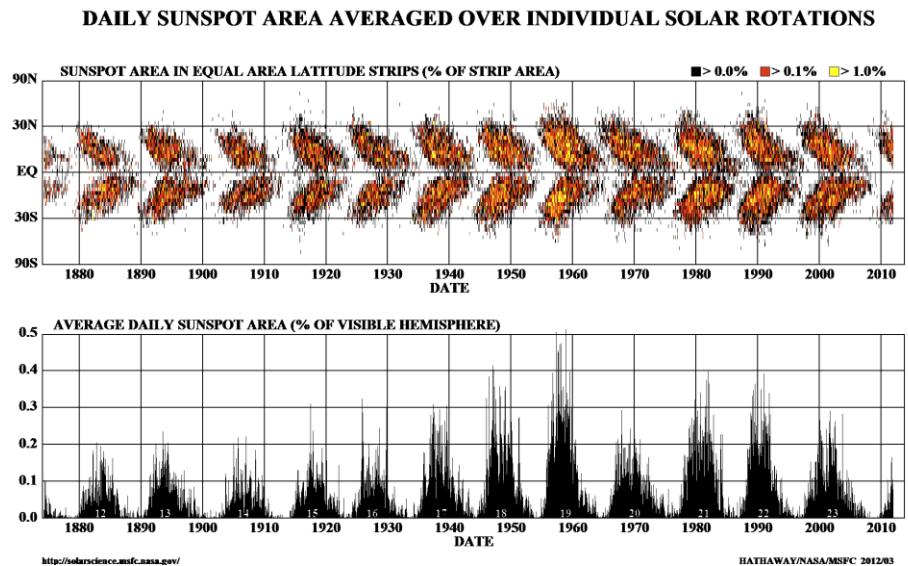
Solar Cycle and Miscellaneous Variable Solar Activity

A brief summary follows, see also my other summary.¹⁷¹

Sunspots

Sunspots, first observed by Galileo, follow a ~11-year cycle from peak activity (measured by *sunspot number*¹⁷²) to peak activity. They tend to arise first at higher latitudes ($\pm 30\text{-}40^\circ$) and gradually extend toward (but not on) the equator in the same cyclical periodicity, producing the familiar Butterfly diagram (upper right, here from NASA¹⁷³):

The temperatures of sunspot *umbrae* (the darkest central parts, as large as 30000 km) are 3,900 - 4,100 K, whereas the surrounding filamentary *penumbrae* are somewhat warmer. Compared to



¹⁶⁶ <http://solarscience.msfc.nasa.gov/dynamo.shtml>

¹⁶⁷ http://en.wikipedia.org/wiki/Magnetic_reconnection

¹⁶⁸ <http://solarscience.msfc.nasa.gov/SolarWind.shtml>

¹⁶⁹ http://www.mcgoodwin.net/pages/spacephysics_ess471.pdf

¹⁷⁰ IMA2 p. 380 and http://en.wikipedia.org/wiki/Stellar_rotation

¹⁷¹ http://www.mcgoodwin.net/pages/spacephysics_ess471.pdf

¹⁷² Boulder and International Sunspot Numbers: <http://spaceweather.com/glossary/sunspotnumber.html>

¹⁷³ NASA image from <http://solarscience.msfc.nasa.gov/images/bfly.gif>

the surrounding photosphere at about 5,777 K, the cooler sunspots have bolometric flux of about 1/5, and are visible as dark spots. Despite their apparent darkness, even the umbra would be quite bright if viewed without a darkening filter.

Using spectral Zeeman splitting (such as with the Fe I 525.02 nm line) and polarization data, the strongest magnetic fields are observed in the umbras, where they are vertically oriented) whereas in the penumbras are found somewhat weaker and more obliquely or horizontally oriented fields. The extra magnetic pressure in a sunspot keeps it from sinking in (IMA2 p. 385).

Sunspots often occur in groups. The lead sunspot of a group has the same magnetic polarity in a solar hemisphere during an 11-year sunspot cycle. In the other hemisphere or during the next 11-year cycle in the same hemisphere, the polarity is reversed. The overall Sun polarity reverses during the sunspot minimum, so that a return to the same polarity and peak sunspot activity takes about 22 years.

The cycling of sunspot activity is accompanied by cyclical variation in the overall solar luminosity of about 0.1% (IMA2 p. 383 and *here*¹⁷⁴). It is hypothesized that frozen-in local magnetic fields associated with sunspots inhibit convection bubbles at the sunspot leading to local cooling. (Note however that somewhat paradoxically, the overall solar luminosity is maximal when sunspots are maximal).

The *Maunder Minimum* (c. 1645 - 1715) was a prolonged period of quiescent solar activity and sunspots, colder Earth temperatures ("Little Ice Age"), and presumably lower solar luminosity. (Evidence to that effect is deduced from carbon-14 and beryllium-10 abundances, isotopes which are affected by cosmic rays, in turn affected by the Sun's magnetic field).¹⁷⁵ John Eddy, who popularized the Maunder Minimum, suggests there may in fact be a longer period variation in solar activity and luminosity that may last for centuries.

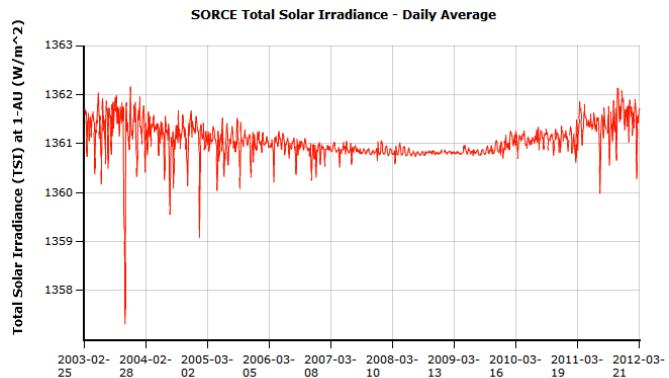
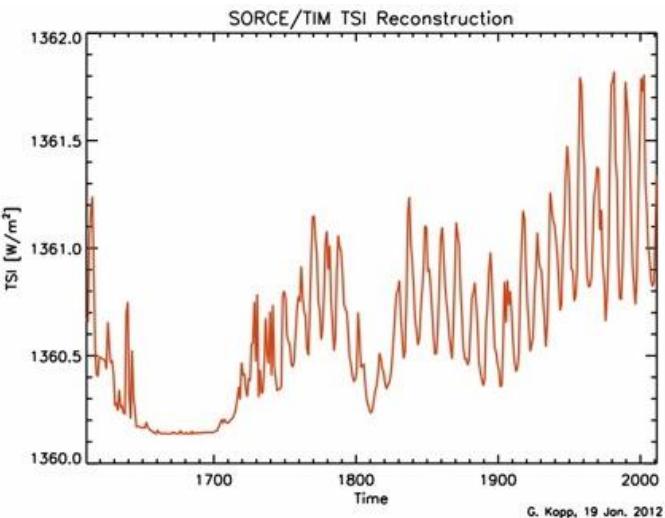
Total Solar Irradiance TSI

TSI is the spatially and spectrally integrated solar radiation incident at the top of the Earth's atmosphere (i.e., "extraterrestrial" and therefore unaffected by Earth's atmosphere). It is usually averaged over at least a day to avoid the transient effects of sunspots, etc., and is often expressed adjusted to 1 A.U. distance from the Sun. The TSI used to be called the "Solar Constant" but it is obviously not constant:

Historical Reconstruction of Total Solar Irradiance

1610 – 2011 (graph to right):

Based on Wang, Lean, and Sheeley, *The Astrophysical Journal*, 625:522, 2005 May 20, and extended with SORCE data... Data for this reconstruction was computed by G. Kopp using TIM V.12 data Jan 19, 2012.¹⁷⁶ SORCE = Solar Radiation and Climate Experiment, a NASA-sponsored satellite mission orbiting at 645 km. (See also above how the luminosity of the Sun has risen over its much longer entire 4.57 Ga history from an initial 0.677 L_☉)



SORCE Daily-Average TSI 2003 – 2012 (graph to right)
Data from Laboratory for Atmospheric and Space Physics (LASP)¹⁷⁷, adjusted to 1 A.U.

¹⁷⁴ http://en.wikipedia.org/wiki/Solar_variation

¹⁷⁵ http://en.wikipedia.org/wiki/Maunder_Minimum

¹⁷⁶ http://lasp.colorado.edu/sorce/data/tsi_data.htm

¹⁷⁷ http://lasp.colorado.edu/lisird/sorce/sorce_tsi/index.html

Using their reported daily average data, the average value of TSI for the period shown (2/25/2003 – 3/25/2012) is 1361.05 W m^{-2} at the upper limit of the Earth's atmosphere adjusted to a distance of 1 A.U. The SORCE average true Earth irradiance (i.e., still extraterrestrial but not corrected to a distance of 1 A.U.) for the same period is nearly the same, namely 1361.24 W m^{-2} .

Other Phenomena

Many of the variable solar phenomena are mentioned in connection with the atmospheric layers.

Comet tails extend beyond the *coma* surrounding the *nucleus* and become visible as comets near the Sun and are illuminated and heated by it, inducing vaporization. The dust reflects sunlight directly and the gases glow from ionization. The streams of dust and gas each form their own distinct tail component. The tail of dust is left behind in the comet's orbit in such a manner that it often forms a curving tail called the *antitail*. The ion tail, made of gases and sometimes called the *true tail*, always points directly away from the Sun. The ionized gas is more strongly affected by the solar wind than is dust, and follows the Sun's magnetic field lines rather than an orbital trajectory. From the Earth, the tails may appear to point in opposite directions.¹⁷⁸

Similar Behavior in Other Stars

Other stars also exhibit magnetic activity cycles (including Zeeman splitting), and slowing of overall rotation. Flare stars, which are main-sequence spectral class M, exhibit rapid changes in brightness possibly due to flares. Star spots have been inferred by their effect on luminosity, especially notable in RS Canum Venaticorum and BY Draconis stars (*IMA2* p. 393 and these¹⁷⁹).

¹⁷⁸ Comet Tails:

- modified from http://en.wikipedia.org/wiki/Comet_tail
- <http://www.universetoday.com/40820/comet-tail/>

¹⁷⁹ Starspots:

- <http://www.livingreviews.org/lrsp-2005-8>
- <http://go.owu.edu/~physics/StudentResearch/2003/BethCademartori/index.html>

The Interstellar Medium; Protostar and Early Star Formation (Chapter 12)

Interstellar Dust and Gas

Star evolution is a cyclical process, in which stars are born from gas and dust in the *interstellar medium* (ISM). Stellar winds, planetary nebulae, and explosive supernovae etc. return some of this material to the ISM. We must consider turbulent gas motions, shocks, magnetic fields, magnetohydrodynamics, thermodynamics, quantum mechanics, chemistry, radiative transfer of the ISM to properly study star evolution.

Dust forms probably mostly from coagulation within ISM molecular clouds, but also from supernova explosions and stellar winds and in the envelopes of cool stars (IMA2 p. 411).

Dust in the ISM contributes to extinction/attenuation, reddening of transmitted light, reprocessing of UV and optical light to IR, scattering of light, and locking up of metals (ZXI lecture 2). These topics are discussed below or elsewhere in this document.

Interstellar Light Extinction (Attenuation) A_λ

This is an astronomical term describing the effects of absorption and scattering of electromagnetic radiation by matter (dust and gas) between an emitting astronomical object and the observer.¹⁸⁰ It is often expressed as a positive magnitude of wavelength-dependent extinction A_λ where apparent magnitude is given by:

$$m_\lambda = M_\lambda + 5 \log_{10}(d) - 5 + A_\lambda$$

The extinction measured in magnitude changes in the visual band and occurring from dust measured is expressed as A_V .

Extinction may be said to account for the dark bands in the Milky Way, the Horsehead Nebula, and the Bok globule B68.¹⁸¹

For wavelength-dependent optical depth defined in $\frac{I_\lambda}{I_{\lambda,0}} = e^{-\tau_\lambda}$, the authors derive $A_\lambda = 1.086$, so that the change in magnitude due to extinction is approximately equal to the optical depth τ_λ . Moreover, $\tau_\lambda = \sigma_\lambda N_d$, where σ_λ = scattering cross section for the dust present and N_d = dust grain column density (number m⁻²) along the line of sight from object to observer. (IMA2 p. 401)

Mie Scattering compared to Rayleigh Scattering

In formulating *Mie Scattering* theory (Gustav Mie, 1868 - 1957), Mie assumed dust grains are spherical of radius a and defined an extinction coefficient $Q_\lambda \equiv \sigma_\lambda / \sigma_g$. His formula applies to light having a wavelength in the same size range as the dust grain diameters ($a \approx \lambda$):

$$\sigma_\lambda \propto a^3 / \lambda$$

Thus, the Mie scattering cross section or probability for $a \approx \lambda$ is greater with shorter than longer wavelengths. When viewing a nebula with scattered light, the observer often sees bluer light from scattering.

For Rayleigh scattering the scattering particles are much smaller, having $a \ll \lambda$, and there is greater Rayleigh scattering of shorter wavelength light. Rayleigh scattering exhibits a strong wavelength dependence, namely $\sigma_\lambda \propto \lambda^{-4}$. This causes, for instance the predominance of scattered blue light in the sky as seen from the ground.¹⁸²

¹⁸⁰ http://en.wikipedia.org/wiki/Extinction_%28astronomy%29

¹⁸¹ Bok globules:

- <http://www.eso.org/public/images/eso0102c/>

- <http://iopscience.iop.org/1538-3881/124/5/2749/fulltext/202251.text.html>

¹⁸² http://en.wikipedia.org/wiki/Rayleigh_scattering

Mie scattering (still with $a \approx \lambda$) is lower for longer wavelengths. Mie scattering accounts for *interstellar reddening*, the reddening of starlight that occurs from passage through dust on its way to the observer. (The amount of this effect can be differentiated from Doppler redshift by analyzing emission and absorption lines.)

For Rayleigh scattering, the lesser scattering of red light results in greater red light transmission in the sky and redder sunsets.

In Mie scattering theory, when λ is large compared to a , Q_λ goes to zero (i.e., small dust particles do not affect long wavelength radiation). When λ is small compared to a , Q_λ becomes constant and $\sigma_\lambda \propto a^2$ (i.e., large dust particles simply block radiation that encounters them).

Molecular Extinction

In actual interstellar extinction measurements, there are regions or “bumps” of greater than expected extinction in the UV range, which is apparently due to other molecules. The peak at 217.5 nm may be attributable to *graphite*, which would therefore be a major component of ISM dust. However, it might also be due to *polycyclic aromatic hydrocarbons* PAH (such as anthracene, coronene, etc.) There is also evidence of IR extinction at 9.7 μ m and 18 μ m from SiO.

Light scattered by dust tends to be slightly polarized by a few percent, and a preferred orientation suggests the presence of a magnetic field in the ISM.

It is likely that interstellar dust partly consists of graphite and/or SiO.

Hydrogen as Dominant Component of the ISM

Hydrogen is the dominant component of the ISM, and is found in H₂ (molecular), H I (neutral H) and H II forms. It is about 70% of the ISM by mass, with He the bulk of the remainder, and much smaller fractions of metals, carbon, silicon, etc.

Hydrogen 21 cm Emission line

The hydrogen 21 centimeter line or H I line, one of the great gifts of nature for radio astronomers, refers to the electromagnetic radiation spectral line that is created by a change in the energy state of neutral hydrogen atoms, H I. The photons emitted have a precise frequency of 1420.40575 MHz (a microwave-radio emission), which is equivalent to a wavelength of 21.10611405 cm in free space. The spectral flux density of 21-cm radiation, like other radio, microwave, and IR emissions, is often expressed in **Jansky** units (1 Jy = 10⁻²⁶ W m⁻² Hz⁻¹, where the Hz⁻¹ allows inclusion of detector bandwidth).

The emission arises as a forbidden transition, therefore improbable, and occurs with a long half-life of 2.9x10⁻¹⁵ s⁻¹. (Thus it can be observed arising only from tenuous gases such as in the ISM, for which the probability of atomic collisions is relatively low.) The H I atom in the ground states may be in one of two electron spin states, that with $m_s = +1/2$ (for which electron and proton have parallel but opposite *anti-parallel* orientations in a magnetic field) and the lower energy state with $m_s = -1/2$ (for which electron and proton have the same and *parallel* orientations in a magnetic field). The flip of the spin to anti-parallel is associated with the photon emission.

These 21-cm photons can penetrate large clouds of interstellar cosmic dust that are opaque to visible light. The center of the 21-cm line remains optically thin over large interstellar distances, so that

$$\tau_H = 5.2 \times 10^{-23} \frac{N_H}{T\Delta v}$$

where N_H is the column density of H I (in atoms m⁻²) and Δv reflects the line broadening from Doppler effects, typically ~10 km s⁻¹... N_H is generally proportional to N_d when A_V is < 1, and it is inferred that dust and gas are distributed together throughout the ISM... This relationship does not apply for denser dust (IMA2 p. 406)

Optically thick dust clouds shield hydrogen from sources of UV, allowing molecular H₂ to persist. Dust can facilitate the chemical combination of H atoms to H₂ (and other molecules such as H₂O) because it provides

fixed combining sites and can absorb some of the binding energy released. Molecular clouds containing H₂ tend to be surrounded and shielded by H I shells.

The 21-cm emission is used in mapping the location of H I in galaxies and the ISM, in measuring radial velocities with Doppler, and estimating magnetic fields with Zeeman effects. They are also used to assess structure and kinematics of galaxies, including the Milky Way.

Detection of Molecular H₂

The H₂ molecule does not emit the 21-cm line and is quite difficult to directly detect at the relatively cool temperatures of the ISM. Its presence is usually inferred based on the amount of a proxy or tracer molecule with more favorable emissions such as CO, which usually has a number abundance of 10⁻⁴ that of H₂ and an emissions at 2.6 mm and 1.3 mm. Other useful molecular and radical tracers are OH (hydroxyl radical, observed at 18-cm and other wavelengths), CS (carbon monosulfide), CH (methyldyne radical), C₃H₂ (cyclopropenylidene), HCO⁺ (formyl cation), and N₂H⁺ (diazonium).

It may also be helpful to study specific isotopes of certain of the atoms making up otherwise identical molecules. The textbook refers to these tracers as *isotopomers*, although it appears that *isotopologue* would be the preferred term.¹⁸³

Interstellar Clouds Classification

A broad and here abbreviated classification is as follows (IMA2 p. 408):

Diffuse H I Clouds: Diffuse H I clouds have temperatures of 30 to 80 K, number densities of 1 to 8 x 10⁸ m⁻³, and masses in the range of 1 to 100 M_⊙.

Diffuse Molecular Clouds (Translucent molecular clouds): These have primarily atomic H (H I) and extinction is 1 < A_V < 5. Molecular H (H₂) can be observed in areas of higher column density. These are like H I clouds but with somewhat higher masses. They have number densities of 5 to 50 x 10⁸ m⁻³. and T of 15 - 50K. The molecular gas is clumpy or filamentary on small scales, helping to explain why stars form in clusters (ZXI lecture 2 and *here*¹⁸⁴).

Giant molecular clouds (GMCs): These are enormous complexes of dust and gas with temperatures of ~15 K, number densities of 1 to 3 x 10⁸ m⁻³, and masses typically 10⁵ to up to 6 x 10⁶ M_⊙, typically about 50 pc across. They are clumpy with local regions of higher density. They may contain *Dark cloud complexes* (like the Horsehead Nebula), and smaller denser clumps. The densest cores may have masses of 10 M_⊙, and *hot cores* may form with masses of 10 - 3000 M_⊙ and T of 100 - 300K. Based on IR observations, hot cores probably harbor embedded massive young O and B stars. They are the breeding grounds of protostars.

Bok globules: These very dark structures typically located in H I regions outside larger molecular complexes, have very high A_V ~ 10, low T ~ 10 K, large number densities n > 10¹⁰ m⁻³, low masses ~ 2 - 50 M_⊙ and small size < 1 pc.¹⁸⁵ They may be dense cores that have been stripped of their surrounding molecular gas by hot massive stars. (IMA2 p. 409)

¹⁸³ Isotopologues and Isotopomers:

• “Isotopologues are molecules that differ only in their isotopic composition. Simply, the isotopologue of a chemical species has at least one atom with a different number of neutrons than the parent.”

<http://en.wikipedia.org/wiki/Isotopologue>

• “Isotopomer: Isomers having the same number of each isotopic atom but differing in their positions. The term is a contraction of ‘isotopic isomer’. Isotopomers can be either constitutional isomers (e.g. CH₂DCH=O and CH₃CD=O) or isotopic stereoisomers [e.g. (R)- and (S)-CH₃CHDOH or (Z)- and (E)-CH₃CH=CHD]”

<http://goldbook.iupac.org/I03352.html>

¹⁸⁴ http://en.wikipedia.org/wiki/H_I_region

¹⁸⁵ http://en.wikipedia.org/wiki/Bok_globule

Interstellar Chemistry

Summarization omitted (see *IMA2* p. 409).

Heating and Cooling of the ISM

Heating in the ISM and ISM clouds arises from cosmic rays, namely charged particles, through the ionization of hydrogen atoms and molecules. The ejected electrons interact with the ISM to increase kinetic energy and T. Heating also arises from ionization of C atoms by UV from starlight, photoelectric ejection of electrons from dust grains by UV, and ionization of H by stellar X-rays, also in some cases shocks from supernovas.

Cooling arises mostly from emission of infrared photons by atoms or molecules previously excited by collisions, etc. These IR emissions typically can pass easily through the cloud and the ISM.

Protostars

Definition of Protostar

Protostars are pre-nuclear burning objects formed from interstellar molecular clouds (*IMA2* p. 412).

According to *IMA2* p. 424, the *protostar* stage begins when the collapsing cloud or object shrinks to a radius of about 5 AU, light can no longer readily escape, and collapse is no longer isothermal but has become adiabatic (so that the core is rising in temperature). See below regarding "When a Protostar becomes a Star".

Timescale for Protostar Evolution

Protostellar evolution proceeds at a slower and more prolonged rate, given by the Kelvin-Helmholtz timescale t_{KH} (see earlier discussion, about 10^7 years for the Sun), rather than by the free-fall timescale t_{ff} (on the order of 10^5 years, thus $t_{KH} \gg t_{ff}$, see *IMA2* p. 416, 425).

The Jeans Criteria

Under what circumstances do protostars form from collapsing cloud globules and cores? In the theory formulated by Sir James Jeans (1877 - 1946), he used the virial theorem and derived that if kinetic energy is insufficient, the cloud will collapse. The condition for collapse of a cloud of mass M_C is

$$M_C > M_J \cong \left(\frac{5kT}{G\mu m_H} \right)^{3/2} \left(\frac{3}{4\pi\rho_0} \right)^{1/2}$$

where M_J is the **Jeans' Mass**, the minimal mass necessary to initiate the spontaneous collapse of the cloud, μ is the mean molecular weight, m_H is the mass of a hydrogen atom = 1.67×10^{-27} kg, and ρ_0 is the initial cloud density. The Jeans criterion for collapse may also be expressed in terms of the minimal cloud radius R_C for which a cloud of density ρ_0 will collapse,

$$R_C > R_J \cong \left(\frac{15kT}{4\pi G\mu m_H \rho_0} \right)^{1/2}$$

where R_J is the **Jeans' Length**.

Bonnor-Ebert Mass

Another important factor that must be considered is an external pressure from the ISM surrounding the cloud, such as an encompassing GMC around a dense core. With the addition of such a surrounding pressure P_0 , the *Bonnor-Ebert Mass* criterion for collapse is

$$M_C > M_{BE} \cong \frac{c_{BE} v_T^4}{P_0^{1/2} G^{3/2}}$$

where

$$v_T \equiv (kT/\mu m_H)^{1/2}$$

and $c_{BE} \approx 1.18$. Diffuse clouds are stable against collapse, but cores of GMCs are not.

Timescale for homologous collapse

In the case of a cloud to collapse satisfying the Jeans criterion and having uniform initial density (*homologous collapse*), a calculation shows that the free-fall timescale t_{ff} for the collapse of the cloud is

$$t_{ff} = \left(\frac{3\pi}{32} \frac{1}{G\rho_0} \right)^{1/2}$$

In this case, all parts of the cloud take the same time to collapse and the free-fall time is independent of initial radius. However, collapse in the case with a starting higher center density is also a possibility and will have a shorter t_{ff} . Sample t_{ff} calculations yield 100,000 – 400,000 years. The rate of collapse steadily accelerates over time.

Fragmentation of Collapsing Clouds

Clouds that are collapsing do not generally collapse into one giant protostar. Rather, the collapsing cloud tends to fragment and preferentially form protostars in groups, a process termed *fragmentation*. Density of the cloud fragments are increasing by many order of magnitude over surrounding regions. Local inhomogeneities include regions that individually satisfy the Jeans Criterion, and these begin their own local collapse. Probably only 1% of the collapsing cloud actually goes into star production. At some point, the collapse is no longer isothermal and becomes adiabatic, and the rising temperature of the collapsing gas must be partly offset by release of heat energy in order for collapse to continue. During adiabatic collapse, the minimum value for the Jeans' Mass of fragments that produce separate collapse is:

$$M_{J_{min}} = 0.03 \left(\frac{T^{1/4}}{e^{1/2} \mu^{9/4}} \right) M_\odot$$

where e is an efficiency factor of about 0.1, and μ is the mean molecular weight. The interpretation is that "fragmentation ceases when the segments of the original cloud begin to reach the range of solar mass objects [1 M_\odot]." (IMA2 p. 419) Increasing temperature halts further collapse and we have *protostars*. Other sophisticated calculations yield a value of $\sim 0.01 M_\odot$ as the lower limit at which fragmentation ceases.

Magnetic Fields

With collapse comes conservation of angular momentum, decreasing moment of inertia, and increasing rate of angular rotation, introducing turbulence and probably magnetic fields. Zeeman splitting measurements show a field of 1 to 100 nT in molecular clouds. With further collapse, the field is frozen-in, becomes stronger, and rising magnetic pressure tends to resist further collapse.

The *supercritical mass* M_B of a cloud, in terms of the magnetic field B , is given by

$$M_B \cong 70 M_{\odot} \left(\frac{B}{1 \text{ nT}} \right) \left(\frac{R}{1 \text{ pc}} \right)^2$$

where B is expressed in nT and R of the sphere is in pc. This is the mass above which gravitational collapse overwhelms the resisting magnetic pressure. Below this mass, the core is *subcritical*.

Ambipolar Diffusion

If the magnetic field strength is locally lower, what would have otherwise been a subcritical mass may be supercritical in that location. Neutral species (such as H I) may migrate in a preferred direction, called ambipolar diffusion. This process can last a long time before free-fall collapse begins. (I have not fully digested this topic.) “In astrophysics, ‘ambipolar diffusion’ refers specifically to the decoupling of neutral particles from plasma in the initial stage of star formation. The neutral particles in this case are mostly hydrogen molecules in a cloud that would undergo gravitational collapse if it were not collisionally coupled to the plasma. The plasma is composed of ions (mostly protons) and electrons, which are tied to the interstellar magnetic field and therefore resist collapse. In a molecular cloud where the fractional ionization is very low (one part per million or less), neutral particles only rarely encounter charged particles, and so are not hindered in their collapse into a star.”¹⁸⁶

Numerical Simulation of Protostellar Evolution and the Onset of Stardom

These complex assessments are done by advanced computer models employing MHD equations . (*IMA2* p. 422). Many details are omitted here, models of various types predict varying results, and I have not digested this topic thoroughly.

Although according to *IMA2* p. 424, the *protostar* stage begins when the collapsing cloud or object shrinks to a radius of ~5 AU, this stage is much larger than the stages depicted in the graphs that follow¹⁸⁷. These graphs simulate the evolution of protostars of varying masses, beginning at the lower right.

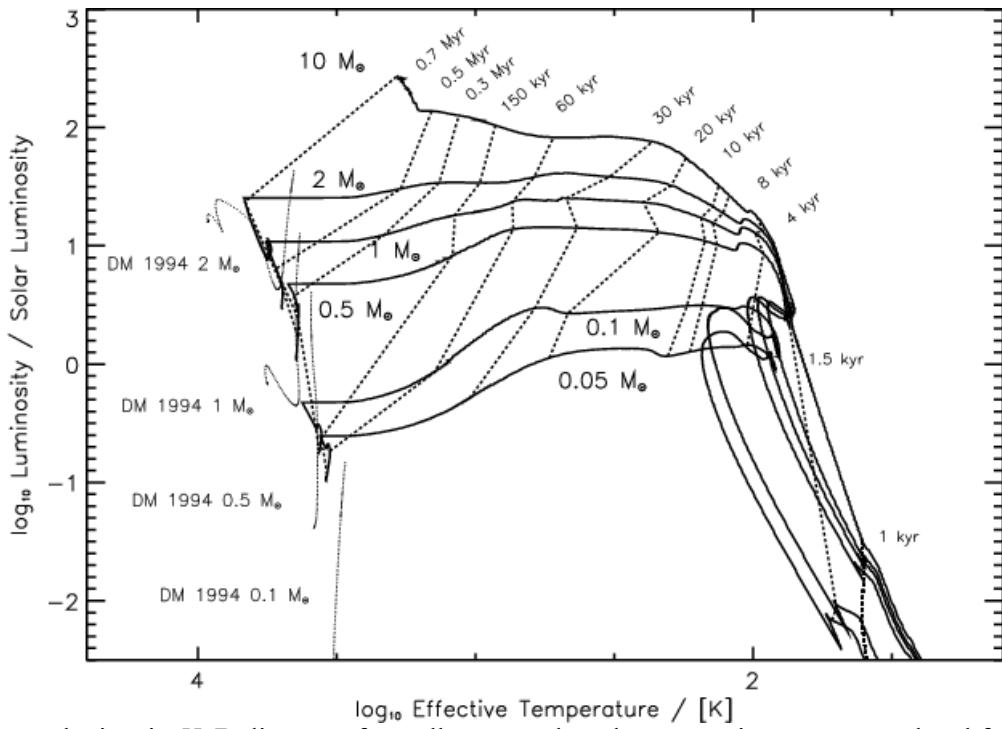
When Does a Protostar Becomes a Star?>: This topic appears to be somewhat hard to pin down. Wuchterl & Tscharnuter 2003 (their data is show in the graphs below) define *stellar zero age* (i.e. when a protostar becomes a star) as follows: “We propose to use as age zero the instant of time, when the interior of the gas spheres is thermally enclosed for the first time. The enclosure means that photons cannot escape from the entire object directly but are radiated from a photosphere. Energy transfer to the photospheric bottleneck then delays the cooling and determines how the thermal reservoir in the interior is emptied... As a practical definition for age zero we propose to use the instant of time when the Rosseland mean optical depth of a gaseous object equals 2/3... For a protostar this instant of time is practically at the end of the isothermal phase of the collapse... It corresponds to the sharp initial luminosity rise (within 1000 yr for 0.05 to 10 M_⊕) that occurs when the first (molecular hydrogen) protostellar cores form.”¹⁸⁸ Some authors simply state the onset of fusion reactions is the point when a protostar becomes a star. Note that the zero stellar age does not occur at the same time as the somewhat later zero-age main sequence (ZAMS).

The following diagram depicts protostar evolution for various masses:

¹⁸⁶ http://en.wikipedia.org/wiki/Ambipolar_diffusion

¹⁸⁷ Wuchterl G & Tscharnuter WM, *A&A* 398, 1081-1090 (2003) , www.aanda.org/articles/aa/full/2003/06/aa8963/aa8963.right.html

¹⁸⁸ *ibid*



Protostellar evolution in H-R diagram for collapse and early pre-main sequence cloud fragments.
Dotted lines are isochrones. See above regarding defining zero stellar age versus ZAMS.
(Wuchterl & Tscharnutter 2003)

As protostars accrete matter, shock waves from infalling material heat the cloud and initially produce its luminosity (i.e., fusion is not yet occurring). The core temperature steadily increases.

Eventually deuterium burning begins in the protostar (at about 10^6 K)¹⁸⁹, because it has a higher reaction cross section than the protium of the P P I chain, and deuterium is the most easily fused nucleus available in protostars. Deuterium burning is strongly T dependent and has the effect of preventing the central T from rising much above 10^6 K. According to Palla and Zinnecker (2002), there are four stages of D burning:¹⁹⁰

- 1) steady state central (core) D burning, fed by infall from convective eddies,
- 2) onset of a radiative barrier around the central region, preventing further infall of matter;
- 3) interior region depleted of deuterium, central deuterium burning stops, allowing the central temperature of the protostar to (?) increase, and
- 4) shell burning of deuterium in expanding convective shells outside the D-depleted central region and radiative barrier.

When central deuterium burning ends, a temporary dip in luminosity and T results, and the star swells temporarily delaying gravitational contraction. After this, T rises, the convecting core becomes radiative, and more steady burning by P P I chain begins.¹⁹¹

The following graph (annotated by MCM from Palla 1999)¹⁹² roughs out the sequence in an accreting protostar from onset of deuterium burning to onset of ZAMS of a main-sequence star. Palla states, “The two curves refer to the evolution computed assuming that matter is accreted directly onto the protostar through a strong shock and to the case where the core surface is treated as a normal photosphere—as in the case of accretion through a circumstellar disk.”

¹⁸⁹ http://en.wikipedia.org/wiki/Deuterium_burning

¹⁹⁰ Palla F; Zinnecker H (2002). *Physics of Star Formation in Galaxies*. Springer-Verlag. pp. 21–22

¹⁹¹ Ibid.

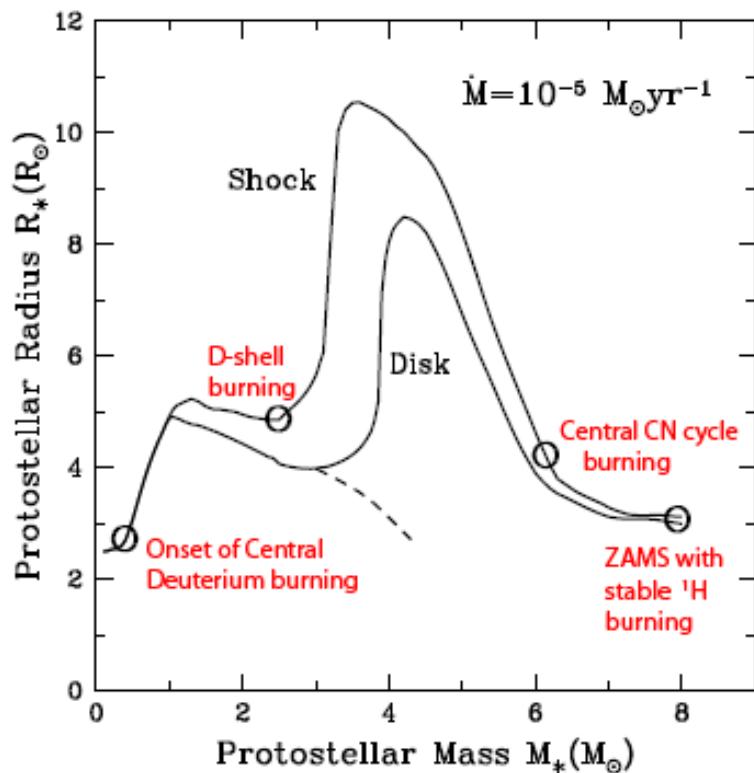
¹⁹² Palla F, The Evolution of Pre-Main-Sequence Stars, 1999.

www.cfa.harvard.edu/events/1999/crete/palla_CreteII.ps

Here, a protostar, starting at a much smaller radius than it had at its inception (i.e., after considerable contraction), accretes from $0 M_{\odot}$ at the start to $8 M_{\odot}$. The open circles and intervals between them show the onset of central deuterium burning, the expansion of the protostar during this burning, onset of Deuterium shell burning with further rapid expansion, depletion of deuterium with contraction, onset of protium 1H burning first with the CNO cycle and eventually with the PP cycle leading at last to the onset of the Zero Age Main Sequence point. See citation above.

The observational search for protostars is underway. They may have a characteristic spectral line profile (IMA2 p. 424): blueshift from infall on the far side, redshift from infall on the near side, and a centrally positioned absorption dip arising from the cloud located between the observer and the center but far out from the region of central collapse.

PRE-MAIN-SEQUENCE EVOLUTION



Pre-Main-Sequence Evolution, Young Stellar Objects, and the ZAMS

Not all models give the same results in this rapidly evolving field.

The *zero-age main sequence* ZAMS is the point in time when a star reaches a state of stable equilibrium in hydrogen (${}^1\text{H}$) burning for which this is the dominant process of energy production, rather than release of gravitational energy, and for which there is hydrostatic equilibrium (IMA2 p. 429 and *here*.¹⁹³) On the H-R diagram, the locus of ZAMS for stars of varying masses is a roughly diagonal line passing from lower right to upper left.

The tracks in the graph to the right (Palla and Stahler 1993) depict ZAMS positions on the H-R diagram, namely at the left end of each track (see the caption).¹⁹⁴ Tracks begin at the *birthline* (complex, see article) of the protostar. Most normal main sequence stars fall slightly up and to the right of their ZAMS position.

The times for stars to reach ZAMS vary inversely with mass: Time to ZAMS is 28,000 yr for a $60 M_{\odot}$ star (virtually no PMS phase at all), 40 Ma for a $1 M_{\odot}$ star , and 68 Ma for a $0.8 M_{\odot}$ star (starting at the Hayashi Track, IMA2 p. 427).

There is a period of time for which a protostar has become a star but is not yet at the ZAMS. This interim stage is termed a *Pre-main-sequence star* (PMS).

A *Pre-main-sequence star* (PMS) is a star that has not yet reached the main sequence. It can be a *T Tauri star* (young, $< 2 M_{\odot}$), an *FU Orionis star* ($< 2 M_{\odot}$), or a *Herbig Ae/Be star* (2 to $8 M_{\odot}$). The energy source of these objects is primarily gravitational contraction (as opposed to hydrogen burning in main-sequence stars)... PMS stars can be differentiated from main-sequence dwarf stars by using stellar spectra to measure the correlation between gravity and temperature. A PMS star will have a larger radius than a main-sequence star, and thus be less dense... Pre-main-sequence stage will last less than 1% of a star's life (in contrast, the star will spend about 80% of its life on the main sequence).¹⁹⁵

Classical calculations showed that a PMS comes to have its outer atmosphere dominated by H^- ion which adds opacity and leads to a convective outer atmosphere. C. Hayashi showed that PMS evolution in this state along the *Hayashi Track* is almost vertical on the H-R diagram (luminosity is decreasing while T_{eff} is only minimally rising).

In the following H-R graphs¹⁹⁶ annotated by MCM, the Hayashi Tracks appears as downstrokes on the far right of each curve. At the lowest point of the Hayashi track, the radiative core allows energy to escape into

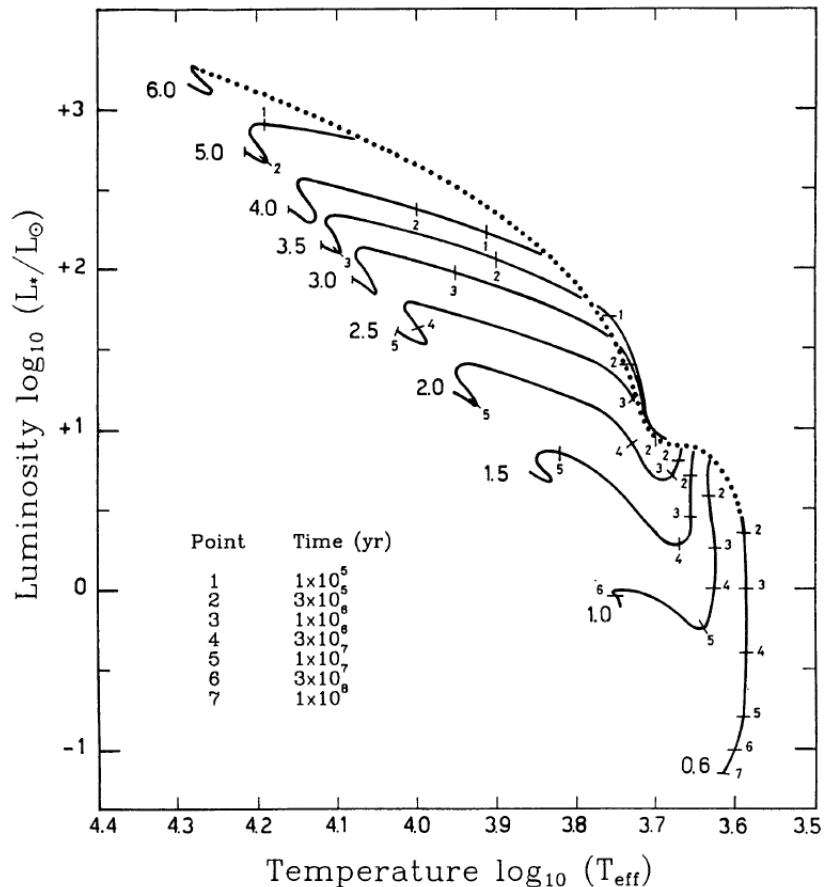


FIG. 10.—Evolutionary tracks in the H-R diagram for all computed masses. We also include the tracks for $M_* = 0.6 M_{\odot}$ and $1.0 M_{\odot}$, as obtained by Parigi (1992). Each track is labeled by the corresponding mass in solar units. Tick marks indicate evolutionary times, as given in the table. Each track starts at the birthline (dotted curve) and ends at the ZAMS. The latter curve has been omitted for clarity.

¹⁹³ http://en.wikipedia.org/wiki/Zero_age_main_sequence

¹⁹⁴ Palla F; Stahler SW, "The Pre-Main-Sequence Evolution of Intermediate-Mass Stars," *Astrophysical Journal* v.418, p.414, 11/1993, <http://adsabs.harvard.edu/doi/10.1086/173402>

¹⁹⁵ http://en.wikipedia.org/wiki/Pre-main-sequence_star

the convective envelope, leading to rising luminosity. At about this minimum point, ^1H nuclear reactions begin in earnest, namely PP I and CNO cycle, while gravitational collapse is providing less of a contribution to L. At the left-hand peak of the luminosity (marked by the short dotted line where the convective core first appears), the central core begins to expand, luminosity and t_{eff} start to decrease. The star thereby approaches the ZAMS state of stable ^1H burning (not explicitly marked but apparently at approximately the left-sided terminus of the tracks).

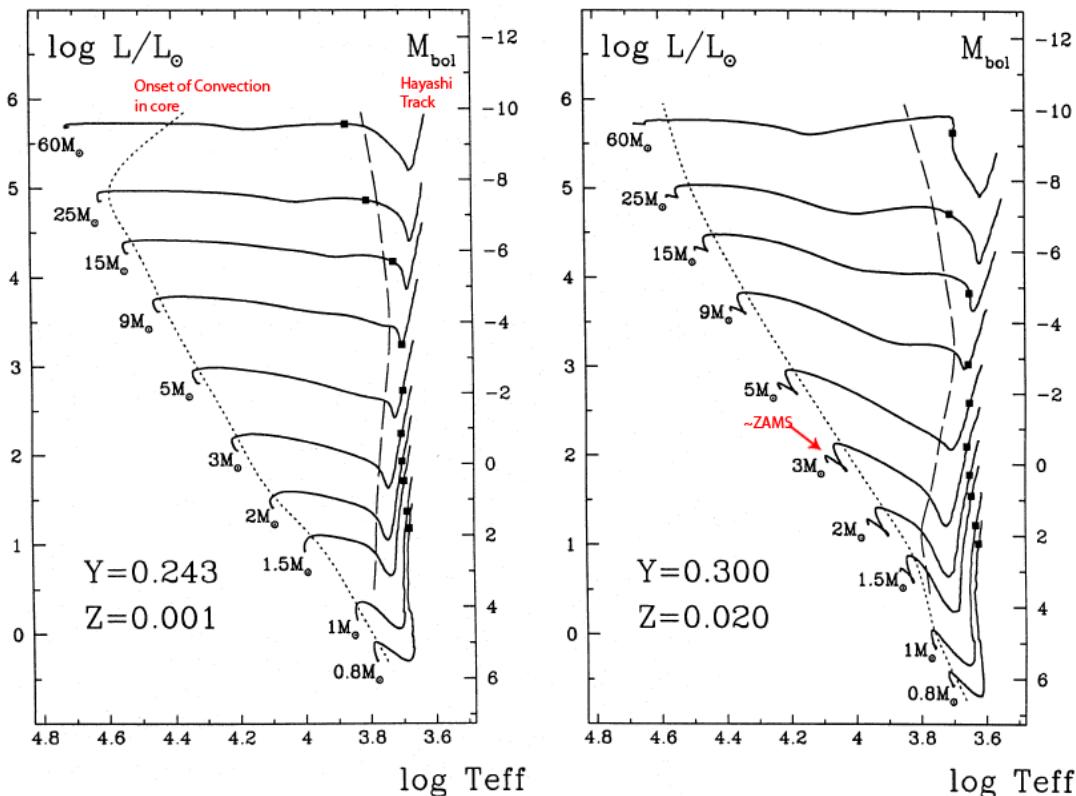
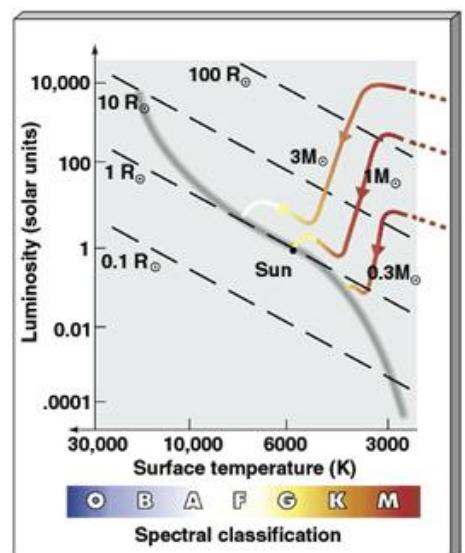


Fig. 5. Theoretical HR diagrams for the computed PMS models at metallicities $Z=0.001$ and $Z=0.020$. Filled squares mark the ignition of deuterium burning. The long dashed line corresponds to the stage where the convective envelope disappears and the stars begin the totally radiative (Hayashi) portion of the tracks. The short dashed line indicates the part of the tracks where the convective core first appears. Near the ZAMS, the effects of the out-equilibrium CNO burning are too small to be observable in the lower metallicity grid.

Optimal calculations should include effects of rotation, magnetic fields, mass loss or accretion, external pressure, feedback caused by infalling matter. Birth lines may be needed.

For lower mass stars ($0.072 - 0.5 \text{ M}_\odot$), the curves are slightly different in that they lack the upstroke just prior to the ZAMS (due to a lack of efficient ^{12}C burning in the core).

The graph to the right is a much simplified depiction of pre-ZAMS evolutionary tracks for stars near 1 M_\odot .¹⁹⁷



¹⁹⁶ Bernasconi PA; Maeder A, "About the absence of a proper zero age main sequence for massive stars", *Astronomy and Astrophysics*, v.307, p.829-839, 1996.

¹⁹⁷ www.castlerock.wednet.edu/HS/stello/Astronomy/TEXT/CHAISSON/BG311/HTML/BG31104.HTM

Massive Stars

For very massive stars, the central temperature quickly becomes high enough to burn ^{12}C and to convert ^1H to ^3He , using the full CNO cycle. The core remains convective even after ZAMS. The current largest star known is the blue hypergiant R136a1, weighing in at 265 M_\odot and with $L = 8.7 \times 10^6 \text{ L}_\odot$. It is a *Wolf-Rayet* star with a surface temperature over 50,000 K.¹⁹⁸

It is possible that stars $> 10 \text{ M}_\odot$ do not form by accretion but by mergers of smaller stars in dense protostellar environments (IMA2 p. 429), but the presence of an accretion disk can reduce the need for mergers.

Brown Dwarfs

Stars with mass $< 0.072 \text{ M}_\odot$ have some nuclear burning, but insufficient to be a main sequence star. Lithium burns at above 0.06 M_\odot and deuterium burns above 0.013 M_\odot but ^1H protium burning is not sustained at mass $< 0.072 \text{ M}_\odot$. For objects $0.013 \text{ M}_\odot < M < 0.072 \text{ M}_\odot$, these are termed *brown dwarfs*. Here, 0.013 M_\odot is about 10 times the mass of Jupiter which is $\approx 9.54 \times 10^{-3} \text{ M}_\odot$. These objects are cool, with spectral types L and T.¹⁹⁹ They are probably common, but are hard to detect due to low luminosity.

Initial Mass Function IMF

The IMF is the probability distribution of mass Φ as a function of stellar mass m [not magnitude], given by

$$dN = \Phi(\log m) d(\log m) = k_1 M^{-p}$$

where dN is the number of stars found in logarithmic mass range from $\log m$ to $\log m + d(\log m)$. It is plotted typically in log-log axes. Massive stars are numerically quite rare, while small stars are abundant. A simple power law is often assumed, such as the exponent $p = 1.35$ utilized by E. E. Salpeter.²⁰⁰

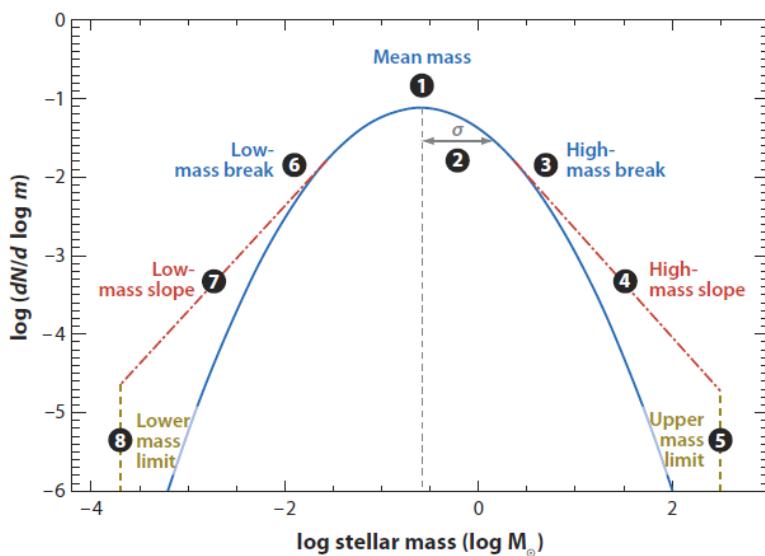


Figure 1

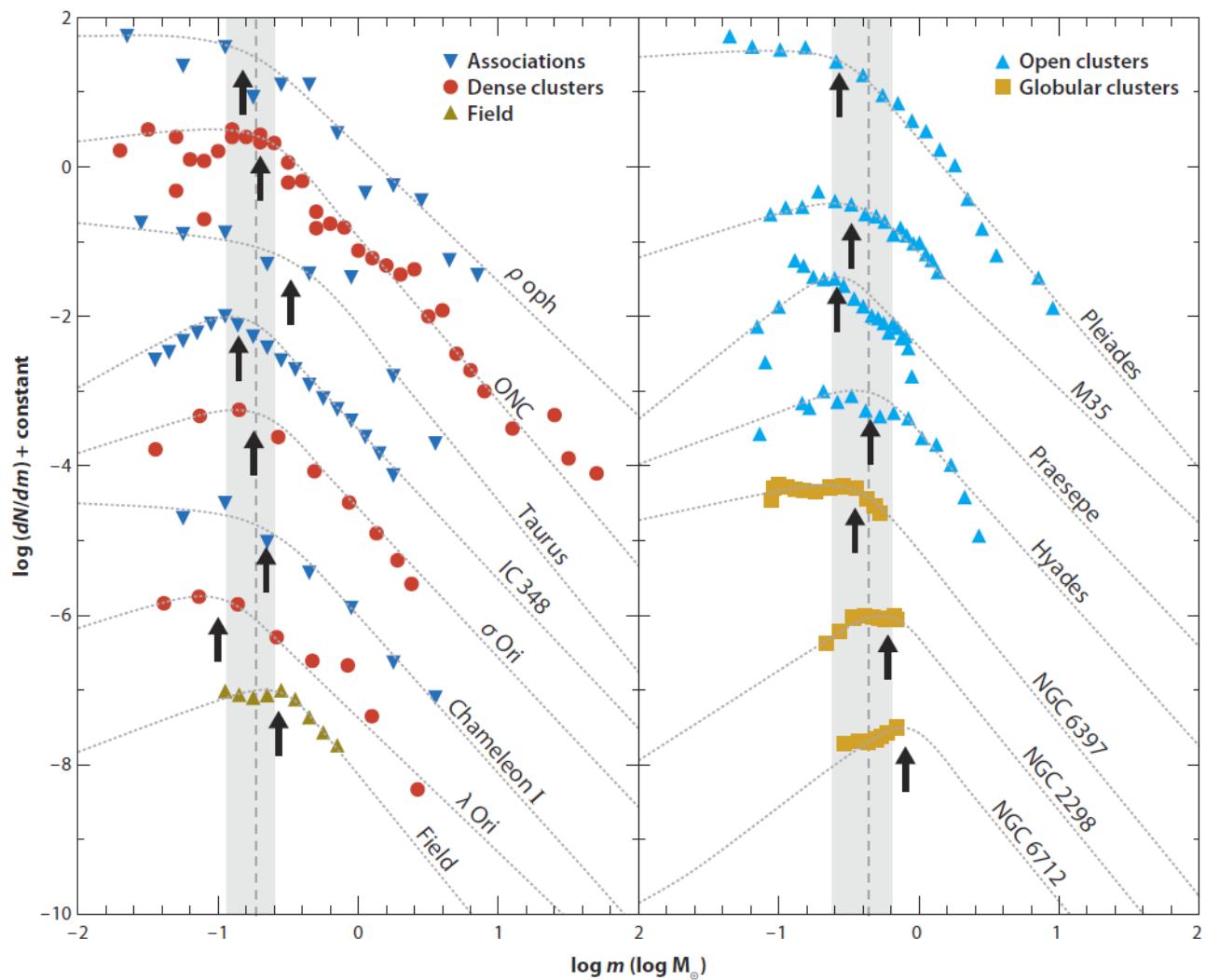
Schematic of an eight-parameter initial mass function (IMF). The “base” of the IMF is approximated as a log-normal distribution (shown as a *solid blue line*) with a (1) characteristic mass and (2) dispersion (σ). On the high-mass side are the additional parameters of the (3) high-mass break, (4) high-mass slope (shown as a *dashed-dotted red line*) and (5) upper mass limit (represented by a *dashed dark yellow line*). Parameters 6, 7, and 8 are the equivalent on the low-mass end.

Schematic of idealized IMF (from a full review of IMF, Bastian et al 2010²⁰¹)

¹⁹⁸ <http://en.wikipedia.org/wiki/R136a1>

¹⁹⁹ http://en.wikipedia.org/wiki/Brown_dwarf

²⁰⁰ http://astro1.physics.utoledo.edu/~megeath/ph6820/lecture12_eqn.pdf



"The derived present-day mass function of a sample of young star-forming regions ..., open clusters spanning a large age range ..., and old globular clusters... Additionally, we show the inferred field star initial mass function (IMF)... The gray dashed lines represent 'tapered power-law' fits to the data... The black arrows show the characteristic mass of each fit (m_p), the dotted line indicates the mean characteristic mass of the clusters in each panel, and the shaded region shows the standard deviation of the characteristic masses in that panel... The observations are consistent with a single underlying IMF, although the scatter at and below the stellar/substellar boundary clearly calls for further study. The shift of the globular clusters characteristic mass to higher masses is expected from considerations of dynamical evolution."

(also from Bastian et al 2010)

"The distribution of stellar masses that form in one star formation event in a given volume of space is called the initial mass function (IMF). The IMF has been estimated from low-mass brown dwarfs to very massive stars. Combining IMF estimates for different populations in which the stars can be observed individually unveils an extraordinary uniformity of the IMF. This general insight appears to hold for populations including present-day star formation in small molecular clouds, rich and dense massive star-clusters forming in giant clouds, through to ancient and metal-poor exotic stellar populations that may be dominated by dark matter. This apparent universality of the IMF is a challenge for star formation theory, because elementary considerations suggest that the IMF ought to systematically vary with star-forming conditions."²⁰²

²⁰¹ Bastian N, et al, "A Universal Stellar Initial Mass Function? A Critical Look at Variations", *Annual Review of Astronomy and Astrophysics* Vol. 48: 339-389, 2010, DOI: 10.1146/annurev-astro-082708-101642.

²⁰² Kroupa P, "The Initial Mass Function of Stars: Evidence for Uniformity in Variable Systems", *Science* 4 January 2002 Vol. 295 no. 5552 pp. 82-91 DOI: 10.1126/science.1067524

H II Regions (Ionized Hydrogen)

Hot O and B stars emit such energetic UV (> 13.6 eV) that surrounding neutral H I gas is ionized to H II. Some of the ions undergo recombinations and de-excitations associated with emission of visible light, in particular the Balmer series H α , causing a beautiful red fluorescence. Examples include the *Orion Nebula M42*. These regions are called *emission nebulae*, some of which are called *planetary nebulae*. The size of an H II region is estimated by the Strömgren radius (*IMA2* p. 432), which can yield values from 0.1 pc to 100 pc.



H II Region in Orion Nebula M42 NGC 1976
associated with the Trapezium cluster and a Giant Molecular Cloud.

This nebula in Orion's sword, south of Orion's belt.

Distance 1500 ly; width 30 arcminutes (4.0 parsecs) square.

Image by NASA/HST, see footnote, 2006 mosaic composite of 5 wavelengths²⁰³.

²⁰³ Orion Nebula:

- Specific H II Region view shown above: <http://www.spacetelescope.org/images/heic0601a/>
- Other views of Orion Nebula: <http://www.spacetelescope.org/images/?search=orion+nebula>

"The Orion Nebula is a picture book of star formation, from the massive, young stars that are shaping the nebula to the pillars of dense gas that may be the homes of budding stars. The bright central region is the home of the four heftiest stars in the nebula. The stars are called the Trapezium because they are arranged in a trapezoid pattern. Ultraviolet light unleashed by these stars is carving a cavity in the nebula and disrupting the growth of hundreds of smaller stars. Located near the Trapezium stars are stars still young enough to have disks of material encircling them. These disks are called *protoplanetary disks* or *proplyds* and are too small to see clearly in this image. The disks are the building blocks of solar systems... The bright glow at upper left is from M43, a small region being shaped by a massive, young star's ultraviolet light. Astronomers call the region a miniature Orion Nebula because only one star is sculpting the landscape. The Orion Nebula has four such stars. Next to M43 are dense, dark pillars of dust and gas that point toward the Trapezium. These pillars are resisting erosion from the Trapezium's intense ultraviolet light. The glowing region on the right reveals arcs and bubbles formed when stellar winds - streams of charged particles ejected from the Trapezium stars - collide with material."²⁰⁴

Large stars, especially O and B star groups called *OB associations*, cause nearby dust clouds to vaporize, molecules to dissociate, and eventually atoms to ionize, creating H II regions from H I regions. High radiation pressure disperses much of the cloud, halting further protostar formation, and possibly even causing loosely bound clusters to become unbound. This is observed in Carina Nebula from nearby star Eta Carinae, and the Eagle Nebula M16 from several massive stars:

Stellar Spire in Eagle Nebula (upper right image): A torrent of ultraviolet light from a band of massive, hot, young stars [off the upper edge of this image] is eroding the pillar of gas and dust (like an eroded desert hoodoo). Distance ~6500 ly, width 6 ly, column height 9.5 ly. (HST composite image)²⁰⁵

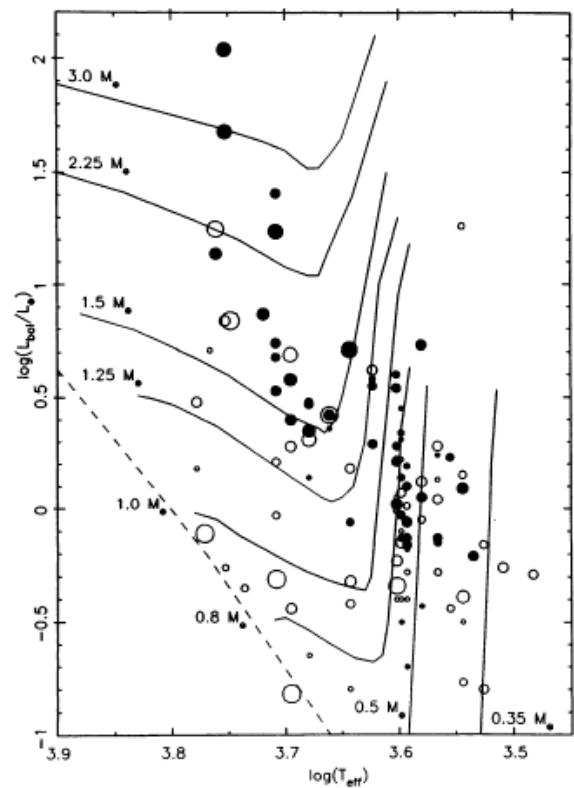


T Tauri Stars

Graphs to right show Evolutionary Tracks for T Tauri stars. Short dashes line = ZAMS.²⁰⁶

These are low-mass pre-main sequence stars (or objects or protostars) representing a transition from stars shrouded in dust (visible only in IR) to main sequence stars. Their masses are 0.5 to about $2 M_{\odot}$. Many emit emission lines including H _{α} , Ca II (H and K), and iron lines, along with lithium absorption. Many emit forbidden lines [O I] and [S I], suggesting very low density.

Inspection of the positions of T Tauri stars in the H-R diagram with stellar evolution models (see graphs above) places them in the region of the Hayashi Track or the trough at the bottom ends of the Hayashi Tracks.²⁰⁷ The ZAMS are found along the dotted line lower left, and the circles represent classical and



²⁰⁴ Text regarding specific view shown above: <http://www.spacetelescope.org/images/heic0601a/>

²⁰⁵ <http://www.spacetelescope.org/images/heic0506b/> 25 April 2005

²⁰⁶ Bertout, Claude, "T Tauri stars - Wild as dust", *Annual review of astronomy and astrophysics*. Volume 27 1989, p. 351-395

²⁰⁷ *ibid.*

“weak-line” T Tauri stars. In other words, T Tauri stars arise and are seen before the star evolves to the ZAMS stage.

The H_a lines of T Tauri stars often exhibit a *P Cygni profile*, suggesting an expanding shell about the object, consisting of:

- a blueshifted absorption trough due to approaching H on the near side of shell, lying directly on the line of sight in front of the star, and which therefore absorbs some of the star’s light
- a broad unshifted emission peak from edges of shell having with no radial movement, and
- a redshifted emission from the part of the receding rear shell that is not obscured by the star.

Their spectra can rapidly change, suggesting these objects are very unstable.

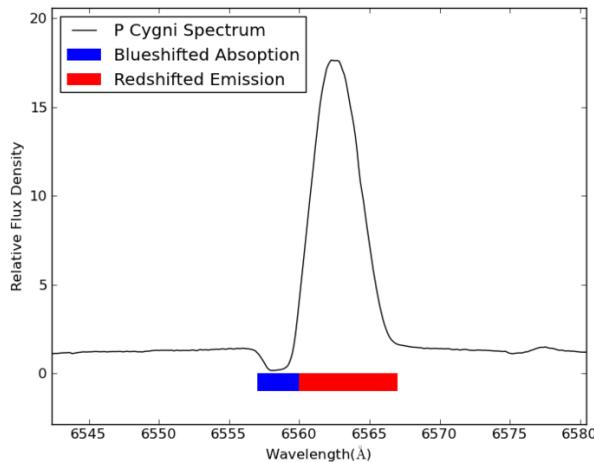


Image to right depicts a P Cygni Profile.²⁰⁸

“T Tauri stars are pre-main sequence stars – the youngest visible F, G, K, M spectral type stars (<2 Solar mass). Their surface temperatures are similar to those of main sequence stars of the same mass, but they are significantly more luminous because their radii are larger. Their central temperatures are too low for hydrogen fusion. Instead, they are powered by gravitational energy released as the stars contract towards the main sequence, which they reach after about 100 million years. They typically rotate with a period between one and twelve days, compared to a month for the Sun, and are very active and variable.”²⁰⁹

FU Orionis Stars

These appear to be T Tauri stars that undergo rapid accretion of mass and marked increase in luminosity, perhaps due to instabilities in a T Tauri accretion disk. (*IMA2* p. 437) They have mass < 2 M_⊙.

Herbig Ae/Be star

A Herbig Ae/Be star (HABe) is a pre-main sequence star of mass 2 to 10 M_⊙, often enveloped in gas and dust. They are “a young (<10Myr) star of spectral types A or B. These stars are still embedded in gas-dust envelopes and may be surrounded by circumstellar disks. Hydrogen and Calcium emission lines are observed in their spectra. They are 2-8 Solar mass objects, still existing in the star formation (gravitational contraction) stage and approaching the main sequence (i.e. they are not burning hydrogen in their center). In the Hertzsprung–Russell diagram these stars are located to the right of the main sequence. They are named after the American astronomer George Herbig (b. 1920), who first distinguished them from other stars in 1960... Sometimes Herbig Ae/Be stars show significant variability.”²¹⁰

Herbig–Haro objects (HH)

These are jets ejected in opposite directions that may be associated with pre-main sequence evolution and young protostars, such as T Tauri stars. The rapidly moving jets excite emission lines on collisions with gas etc. (such as seen with HH47). The jets may be associated with many of the characteristics of protostars, including emission lines, mass loss, jets, and possibly luminosity variations (*IMA2* p. 437).

²⁰⁸ http://en.wikipedia.org/wiki/P_Cygni

²⁰⁹ http://en.wikipedia.org/wiki/T_tauri_star

²¹⁰ http://en.wikipedia.org/wiki/Herbig_Ae/Be_star

"Herbig-Haro objects (HH) are small patches of nebulosity associated with newly born stars, and are formed when gas ejected by young stars collides with clouds of gas and dust nearby at speeds of several hundred kilometers per second. Herbig-Haro objects are ubiquitous in star-forming regions, and several are often seen around a single star, aligned along its rotational axis... HH objects are transient phenomena, lasting not more than a few thousand years. They can evolve visibly over quite short timescales as they move rapidly away from their parent star into the gas clouds in interstellar space (the interstellar medium or ISM). Hubble Space Telescope observations reveal complex evolution of HH objects over a few years, as parts of them fade while others brighten as they collide with clumpy material in the interstellar medium... The first astronomers to study them in detail were George Herbig and Guillermo Haro, after whom they have been named."²¹¹

Young Stars with Circumstellar Disks

Young stars often have these disks, such as Vega²¹² and β-Pictoris²¹³ and HH30²¹⁴. The disks may be *accretion disks*, and clumps of material fall from the disk into the star. In some cases they may be *debris disks* (i.e., disks formed by fragmentation of previously formed objects by collisions). Protoplanets may be present. The central star may be masked by dust (as is the case with HH30).

Circumstellar disks appear to be common during collapse of protostars. This is predicted by the spin-up of the cloud during collapse to conserve angular momentum. The typically slow rotation of main sequence stars might seem inconsistent, but suggests that angular momentum has been transferred away from collapsing protostars and stars that would otherwise be rotating rapidly. This may take place with coupling of magnetic fields in the star's convection zones to the ionized stellar winds, or other mechanisms. (*IMA2* p. 441 and *here*²¹⁵)

Proplyds

These are probably circumstellar disks and appear to be protoplanetary disks associated with young stars of age < 1 million years. The disks have mass much greater than $4 M_{\oplus}$.²¹⁶

²¹¹ http://en.wikipedia.org/wiki/Herbig%E2%80%93Haro_object

²¹² <http://www.solstation.com/stars/vega.htm>

²¹³ <http://www.solstation.com/stars2/beta-pic.htm>

²¹⁴ Watson AM and Stapelfeldt KR, "Asymmetry and Variability in the HH 30 Circumstellar Disk" *The Astronomical Journal* Jan 22, 2007 p. 133 845, <http://iopscience.iop.org/1538-3881/133/3/845>

²¹⁵ Lignières F, Catala C, & Mangeney A., "Angular momentum transfer in pre-main-sequence stars of intermediate mass.", *Astronomy and Astrophysics*, v.314, p.465-476, 1996

<http://articles.adsabs.harvard.edu/full/1996A&26A...314..465L>

²¹⁶ A Megagerie of 30 Proplyds:

<http://www.spacetelescope.org/static/archives/images/screen/heic0917aa.jpg>

Main Sequence and Post-Main-Sequence Stellar Evolution (Chapter 13)

Main sequence stars (MSS) are characterized by their burning of hydrogen to helium in their cores. A star leaves the main sequence when its core is no longer burning hydrogen (*IMA2* p. 457). Here we follow these stars, beginning at the zero-age main sequence (ZAMS) previously defined. This is a very actively studied and rapidly evolving field for which various complex models are not necessarily in agreement. I can't pretend to have the latest or definitive information here.

Evolution of the Main Sequence

Stellar Evolution Timescales

For pre-main sequence (PMS) stars, their timescales are the free-fall timescale t_{ff} and the thermal Kelvin-Helmholtz timescale (t_{KH} , about 10^7 yr for the Sun).

For main sequence stars (MSS), their timescales are also governed by the durations of thermonuclear reactions (about 10^{10} yr for the Sun). This much longer timescale accounts for why 80% of stars in our neighborhood are MSS—this stage simply lasts longer.

Width of the Main Sequence

The main sequence is properly depicted in H-R diagrams as a broad band (see “A modern Hertzsprung–Russell diagram” above and *here*²¹⁷) because the actual distribution of observed stars is broad. The variation is due to differences in chemical composition, stage of evolution, and observational imprecision, etc.

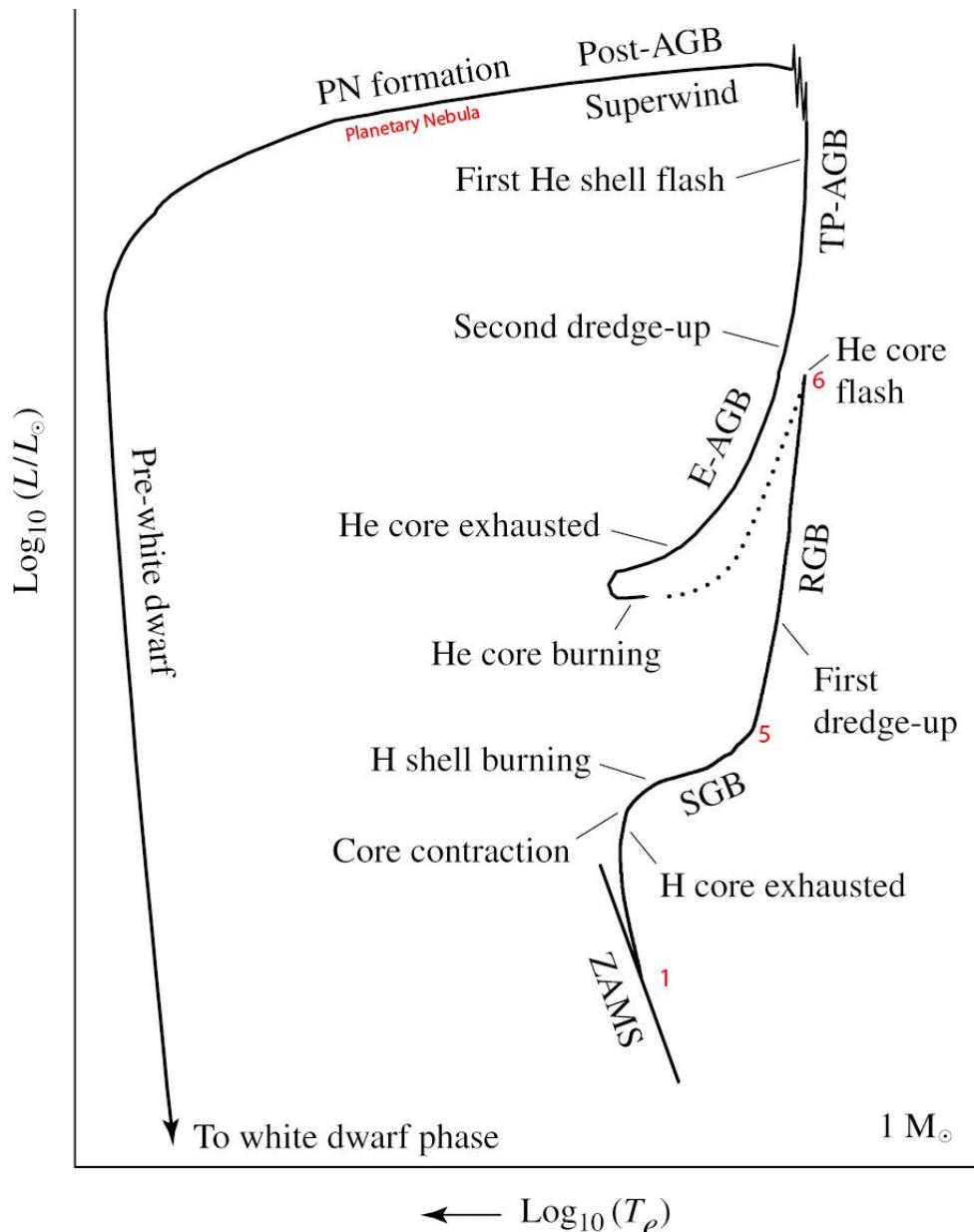
Graphs of Overall Stellar Evolution

The following images are modified from figs. 13.4 and 13.5, *IMA2* p. 458–9, but obtained from another source, annotated by MCM:²¹⁸

²¹⁷ <http://msedbogwaterfilter.blogspot.com/2011/01/review-questions-chapters-15-16-nature.html>

²¹⁸ Images are modified from figs. 13.4 and 13.5, *IMA2* p. 458–9, but obtained from

www.physics.sfsu.edu/~fischer/courses/Astr420/lectures/Ch13.pdf



Evolution of low-mass Sun-like $1 M_\odot$ stars from ZAMS to the final White Dwarf stage

ZAMS = zero-age main sequence (ZAMS);

SGB = SubGiant Branch;

RGB = Red Giant Branch;

E-AGB = Early Asymptotic Giant Branch;

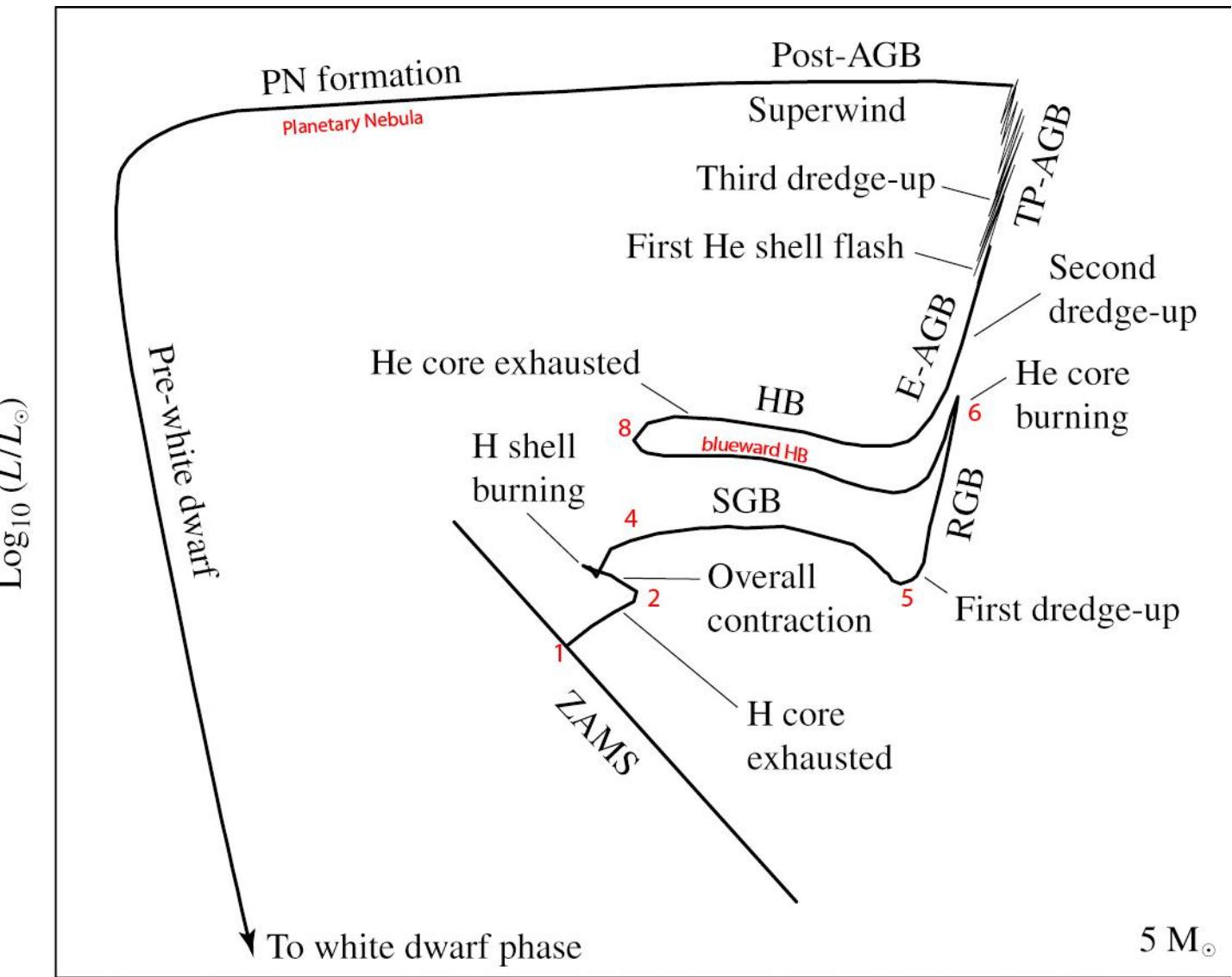
TP-AGB = Thermal Pulse Asymptotic Giant Branch;

Post-AGB = Post-Asymptotic Giant Branch;

PN = Planetary Nebula

(Temp increases toward the left; image annotations in red added by MCM)²¹⁹

²¹⁹ ibid.



$\leftarrow \text{Log}_{10} (T_e)$

Evolution of intermediate-mass $5 M_{\odot}$ stars from ZAMS to the final White Dwarf stage

ZAMS = zero-age main sequence (ZAMS); SGB = SubGiant Branch;

RGB = Red Giant Branch;

HB = Horizontal Branch (blueward and redward portions);

E-AGB = Early Asymptotic Giant Branch;

TP-AGB = Thermal Pulse Asymptotic Giant Branch;

Post-AGB = Post-Asymptotic Giant Branch;

PN = Planetary Nebula

(Temp increases toward the left; image annotations in red added by MCM)²²⁰

Low-Mass Main-Sequence Evolution (including of the Sun) from ZAMS to the End of MSS

These are $0.3 M_{\odot}$ to $1.2 M_{\odot}$. In the upper part of this mass range, stars including the Sun have a radiative core because for much of their time as MSS, they are not emphasizing the temp-dependent CNO cycle (as is the case for stars $> 1.2 M_{\odot}$). The lower mass stars in this range have convective cores.

Rising T, R, and L: Recall that L and R have steadily risen since ZAMS, at least for the Sun (*IMA2* p. 350). The Sun is thought to be 4.57 Ga from ZAMS. The rise in L and R occurs because the fission reactions

²²⁰ ibid.

increase mean molecular weight μ in the core (*IMA2* p. 293). According to the ideal gas law previously presented (*IMA2* p. 291), if density and T were held constant, and μ rises, contraction in the core occurs and gravitational energy is released, increasing interior T . Rising T increases the radius inside of which nuclear reactions can occur. The rate of the PP reaction also increases with rising T . The increased energy output propagates to the surface, and thus the star's luminosity and T_{eff} gradually increase.

The evolutionary tracks of MS stars from 0.8 to $25 M_{\odot}$ are nicely depicted in detail in fig. 13.1 of *IMA2* p. 448, starting at the ZAMS = point 1. The table 13.1 accompanying fig. 13.1 (*IMA2* p.449) show dramatic differences in the burning time for larger vs. smaller stars. The Sun at $1 M_{\odot}$ reaches point 2 in 7048 Ma whereas the $0.8 M_{\odot}$ star reaches point 2 in 18829 Ma and a $25 M_{\odot}$ star reaches point 2 in only 6.3 Ma. (The criteria for point 2 is not stated, but for larger stars it appears to be the point where decline of T_{eff} ceases and T_{eff} begins to rise.) Our Sun is currently on the path between 1 and 2. The interior physical and isotopic structures of the present-day Sun (parameters graphed as a function of radius) are depicted in *IMA2* p.450 (based on Bahcall 2001), and of the Sun near point 3 (about 9.8 Ga from ZAMS) in *IMA2* p. 451 (based on Iben 1967). Comparing these graphs, central depletion of ^1H becomes apparent.

By point 3, nuclear burning continues in a shell surrounding the ^1H depleted core (from which no further luminosity is arising), but overall and surface luminosity continues to rise. The star moves further to the red in the H-R diagram (T_{eff} decreases) as R and L increase, but this is still a MSS.

Mario Schönberg (1914–1990) and Subrahmanyan Chandrasekhar (1910–1995) calculated in 1942 the **Schönberg–Chandrasekhar limit (SC limit)**. This is the maximum fraction of a star's mass, typically 10 to 15%, that can exist as an isothermal burnt out helium core (too cold for triple alpha burning), and for which the core is still able to support the enveloping material above it. “If this limit is exceeded, as can only happen in massive stars, the core collapses, releasing energy that causes the outer layers of the star to expand to become a red giant.”²²¹ (Note: this limit is not the same as the *Chandrasekhar limit*, which is the maximum mass of a stable white dwarf star, about $1.4 M_{\odot}$, beyond which, a neutron star or black hole may form.)

The calculation for SC limit is based on the mean molecular weights of the envelope versus the isothermal core (*IMA2* p. 451):

$$\left(\frac{M_{\text{ic}}}{M}\right)_{\text{SC}} \approx 0.37 \left(\frac{\mu_{\text{env}}}{\mu_{\text{ic}}}\right)^2$$

“The core increases in mass as the shell burns its way outwards through the star. If the star's mass is less than approximately 1.5 solar masses, the core will become degenerate before the Schönberg–Chandrasekhar limit is reached. But if the mass is greater than approximately 6 solar masses, gravitational collapse [on a Kelvin-Helmholtz timescale] will release so much energy that the core will not become isothermal prior to the start of helium fusion. In the remaining case, where the mass is between 1.5 and 6 solar masses, the core will grow until the limit is reached, at which point it will contract rapidly until helium starts to fuse in the core.”²²²

When this SC limit is reached (point 4 in the diagram), a star of $< 1.2 M_{\odot}$ is no longer a MSS (*IMA2* p. 452) and continues to redden with falling T_{eff} . The core acquires partial *degeneracy*, the percent being greater in less massive stars. Degeneracy is a state wherein the electrons occupy the lowest possible energy levels, allowing for the Pauli exclusion principle for fermions, and resist being packed in any tighter. Degeneracy provides Temp-independent pressure resisting collapse, and smaller stars may never reach the SC limit before the next stage of nuclear burning begins, because of degeneracy pressure.

Main-Sequence Evolution of Massive Mass Stars from ZAMS to the End of Main Sequence Star

These Massive Mass Stars are here defined as $> 1.2 M_{\odot}$. Unlike those just below $1.2 M_{\odot}$, these have a convective core which keeps their material well mixed. When a $5 M_{\odot}$ star reaches point 2 (having $X_c = 0.05$, the mass fraction of ^1H), the entire star begins to contract (presumably the definition of point 2). L rises from gravitational energy release, and T_{eff} rises. The graph thus bends up and to the left. For a star of $> 1.2 M_{\odot}$, this point 2 is the end of the MSS phase.

²²¹ http://www.daviddarling.info/encyclopedia/S/Schonberg-Chandrasekhar_limit.html

²²² http://en.wikipedia.org/wiki/Sch%C3%B6nberg%E2%80%93Chandrasekhar_limit

Late Stages of Stellar Evolution (Post-Main Sequence)

A star leaves the main sequence when its core is no longer burning hydrogen and consists mainly of He (IMA2 p. 457). The textbook considers stars of 1 and 5 M_{\odot} mass for detailed evolution descriptions, which I have simplified somewhat.

For a 1 M_{\odot} Post-MS star, the core begins to contract, H-burning takes place in shells outside the core, luminosity increases, R increases slightly, and T_{eff} decreases. The star becomes a red giant and evolves up and to the right on the H-R diagram.

For a 5 M_{\odot} star, the entire star begins to contract, releasing gravitational potential energy. The L rises slightly, R decreases, and T_{eff} rises. Eventually, a shell of H starts to burn, the outer envelope expands, L decreases, T_{eff} drops.

For both 1 M_{\odot} and 5 M_{\odot} stars, the helium core increases in mass from hydrogen shell burning. The envelope expands, T_{eff} drops, and redward evolution occurs across the **Subgiant Branch SGB**. For the 5 M_{\odot} star, there is actually a terminal temporary drop in L (point 5 in the figure 13.1 graph).

The expansion of the stellar envelope and the decline in T_{eff} causes a rise in photospheric opacity from H⁻ ion, causing a convection zone to form near the surface for both sizes, a zone extending deep into the star. The star begins to rise rapidly in L with modest further drop in T_{eff} along the **Red Giant Branch RGB** (a path similar to the Hayashi track of pre-MS stars).

Deepening convection mixes previously unmixed regions, and *lithium burning* occurs and Li becomes widely depleted. The photospheric abundance of Li decreases, and ³He increases. Several other isotopic abundance ratios change (IMA2 p. 461). The observable transport of materials by convection from the deep interior to the surface is termed the **First Dredge-up Stage**. The RGB is a time of hydrogen-burning shell.

Luminosity rises until the **Red Giant Tip** of the RGB phase is reached.

In stars of < 1.8 M_{\odot} but greater than about 0.5 M_{\odot} , the release of triple alpha fusion energy in the core is delayed by core cooling. The core becomes strongly electron degenerate. "The helium [in the core] is degenerate, meaning it is supported against gravity by quantum mechanical pressure rather than thermal pressure. Thus an increase in the temperature in the material undergoing fusion does not act to expand the material and by doing so cool, and there is no regulation of the rate of fusion." ²²³ Eventually the triple alpha fusion reaction begins in the degenerate core, and is very transiently explosive (termed the **Helium Core Flash**, in which the core luminosity is extremely high, up to $10^{11} L_{\odot}$, but lasting for only a few seconds). Much of the energy released is absorbed in the outer layers. This discontinuity makes further modeling for such low mass stars difficult until quiescent He core burning and H shell burning conditions arise.

In larger stars such as for the 5 M_{\odot} star illustrated, the T (1.3×10^8 K) and density are high enough for QM tunneling to overcome the ⁴He nuclear Coulomb barrier, allowing triple alpha He burning fusion processes to begin (generating ¹²C, some of which combine with alphas to yield ¹⁶O). "A star with mass greater than about 2.25 solar masses starts to burn helium without its core becoming degenerate and so does not exhibit this type of helium flash."

In very low mass stars (< ~0.5 M_{\odot}), the core is never hot enough to ignite helium burning. The degenerate helium core will keep on contracting, and finally becomes a **helium white dwarf**. ²²⁴

During He burning, the core and the hydrogen shell expand, cooling it and abruptly decreasing L. The envelope contracts, and T_{eff} rises, and the star moves along the lower **blueward limb** of the **Horizontal Branch HB** (which occurs for both low and intermediate mass stars). This phase is the helium-burning analog of the hydrogen-burning main sequence, but lasts for a much shorter time. The furthest point to the left (hottest T_{eff}) on the HB is attained: before the He has been depleted; the mean molecular weight has increased; the core is contracting; and the envelope is cooling and expanding. The upper and **redward limb** of the **HB** then begins, and soon the core He is depleted. Many stars exhibit instabilities during this phase, exhibiting pulsations in L, R, T, etc.

Asymptotic Giant Branch AGB: When luminosity begins to rise, the **Early Asymptotic Giant Branch EAGB** has been reached. (The AGB is so named for seeming to approach asymptotically to the RGB.) Energy output

²²³ http://en.wikipedia.org/wiki/Helium_flash

²²⁴ Example with M = 0.18-0.19 M_{\odot} : <http://arxiv.org/abs/astro-ph/0404291>

is dominated by the He burning shell just outside the core, the core is mainly C and O with little He remaining, the envelope is expanding, L is rising, T_{eff} is declining somewhat (initially). The convective envelope deepens, a **Second Dredge-up** of interior isotopes occurs, and this brings He and N to the envelope.

The **Thermal-Pulse Asymptotic Giant Branch TP-AGB** ensues, with H-shell burning reignited. The more deeply positioned He-burning shell exhibits quasiperiodic On/Off burning, termed **Helium Shell Flashes**, which are associated with dumping of “helium ash” to the lower layers and abrupt rises in L... Pulses last a few hundred years.²²⁵ The period between pulses ranges from thousands of years for stars of $5 M_{\odot}$ to hundreds of thousands of years for low mass stars of $0.6 M_{\odot}$. Iben has given us a model (see graph to follow) showing for a $0.6 M_{\odot}$ star over 2.4 million years the predicted helium flashes, occurring in similar but not identical intervals of about 130,000 - 325,000 years. A similar model for a larger star is depicted in the graph that follows the first graph:

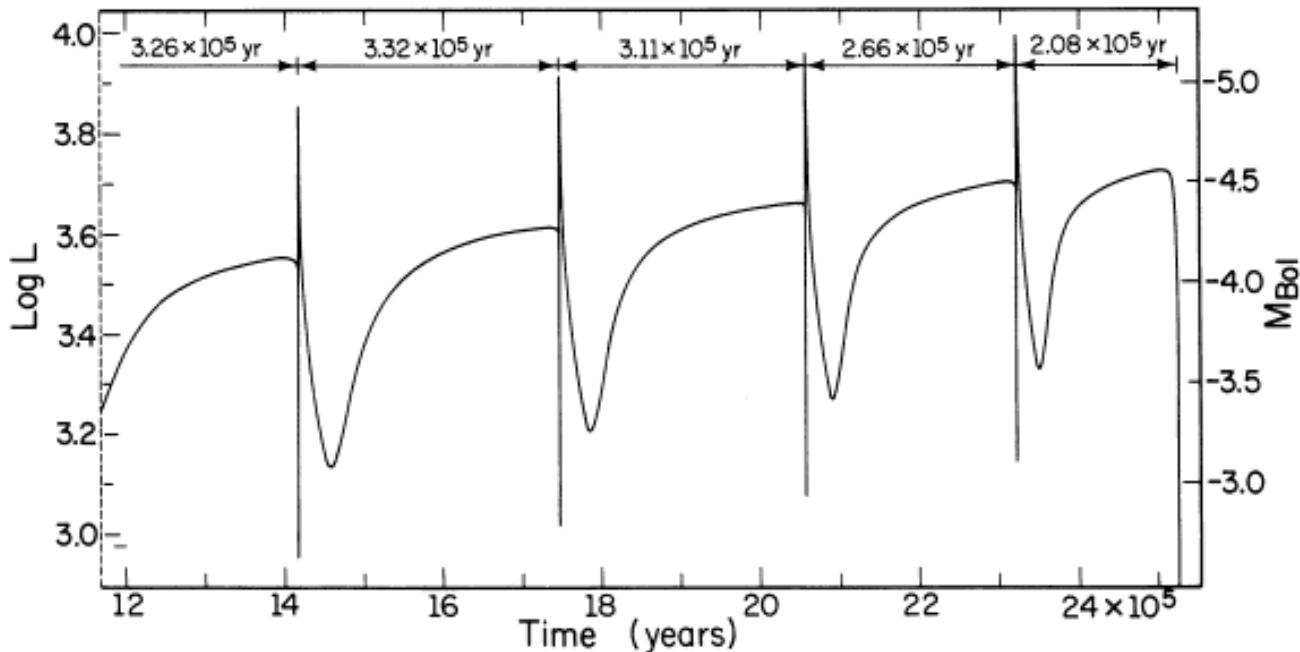


FIG. 2.—Time dependence of luminosity and bolometric magnitude during the thermally pulsing phase for a model of mass $0.6 M_{\odot}$

This graph is extracted from an article by Iben 1982²²⁶ regarding modeling of Helium Shell Flashes on the Thermal-Pulse Asymptotic Giant Branch of a $0.6 M_{\odot}$ star. I have simplified it by deletion of the upper 1/2 of the graph that depicts helium flashes from 0 to 11.5×10^5 yr.

²²⁵ http://en.wikipedia.org/wiki/Helium_flash

²²⁶ Iben, I., Jr., “Low mass asymptotic giant branch evolution. I”, *ApJ, Part 1*, vol. 260, 1982, p. 821-837.

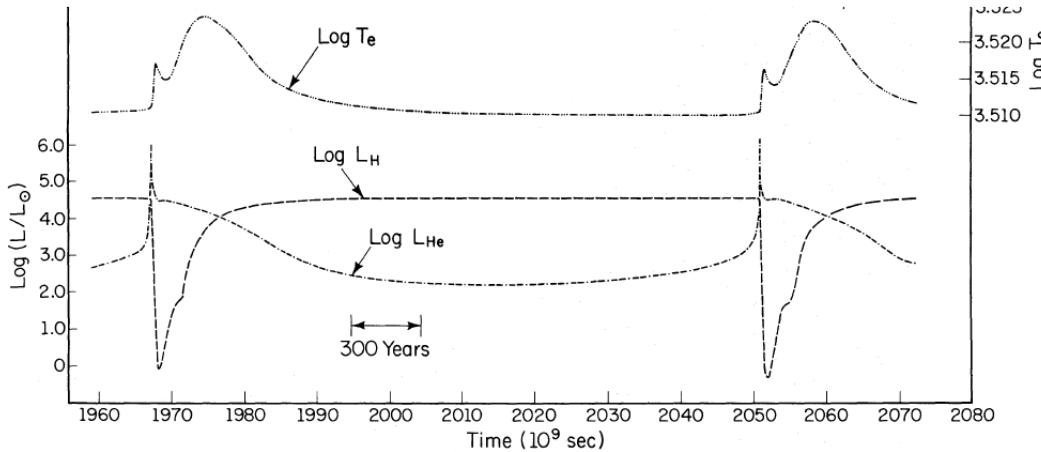


FIG. 1.—Time dependence of several model characteristics during and between the eighth and ninth thermal pulses.
 T_{eff} = surface temperature, L_{H} = power output from hydrogen-burning reactions, L_{He} = power output from helium-burning reactions

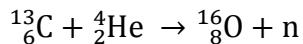
This graph is extracted from Iben's 1975 article²²⁷ regarding modeling of Helium Shell Flashes ("relaxation oscillations") on the Thermal-Pulse Asymptotic Giant Branch of a $7 M_{\odot}$ star. I have simplified it by deleting information on parameters other than luminosity and T_{eff} .

IMA2 p. 465 raises the possibility that certain pulsating **Long Period Variables (LPVs)**, which have periods of 100 to 700 days, including the subclass of **Mira variable stars**, may be experiencing shell flashes that could account for the observable changes in periods, etc.

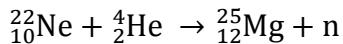
A **Third Dredge-up** may occur in conjunction with the helium flashes. For stars $> 2 M_{\odot}$, the convection zones will extend deep enough to dredge up Carbon. This can cause the spectrum of an oxygen-rich giant star (for which $N_{\text{O}} > N_{\text{C}}$) to transform into that of a **carbon star** (for which $N_{\text{C}} > N_{\text{O}}$). Such carbon stars are designated with spectral class C. There are also spectral class S stars for which ZrO lines are present and in some cases replacing TiO lines.

S-Process Nucleosynthesis: Some TP-AGP stars exhibit ^{99}Tc in their spectra, a radioisotope with half-life of only 211,000 years, indicating it must have relatively recently been dredged up from the deeper sites of nucleosynthesis. (This is not the metastable form of technetium — ^{99m}Tc — which has a much shorter half-life and is used for medical diagnosis.) ^{99}Tc is formed by **s-Process nucleosynthesis** (here, s=slow). In this slow process, a *slow neutron* is captured by the nucleus of a stable isotope that then decays by beta-decay to a stable daughter having 1 greater atomic number than the parent. These decays to new isotopes are able to occur when the density of slow neutrons is sufficiently low, allowing the decay to occur before another neutron can be captured.

This s-process accounts for about 1/2 of the isotopes synthesized [in stars] that are heavier than iron. "The **s-process** is believed to occur mostly in Asymptotic Giant Branch stars. In contrast to the **r-process** [r=rapid] which is believed to occur over time scales of seconds in explosive environments [such as core-collapse supernovae Types Ib, Ic and II], the S-process is believed to occur over time scales of thousands of years. The extent to which the s-process moves up the elements in the chart of isotopes to higher mass numbers is essentially determined by the degree to which the star in question is able to produce neutrons, and by the amount of iron in the star's initial abundance distribution. Iron is the "starting material" (or seed) for this neutron capture - beta-minus decay sequence of synthesizing new elements."²²⁸ Neutrons arise in stellar nuclear reactions primarily by the following reactions,



and



²²⁷ Iben, I., Jr., "Thermal pulses; p-capture, alpha-capture, s-process nucleosynthesis; and convective mixing in a star of intermediate mass", *Astrophysical Journal*, vol. 196, Mar. 1, 1975, pt. 1, p. 525-547., 3/1975.

²²⁸ <http://en.wikipedia.org/wiki/S-process>

(IMA2 p. 313 shows 2 other stellar reactions that yield neutrons.) Because of low neutron fluxes, the s-process is not capable of producing any significant amounts of heavy radioactive isotopes such as thorium or uranium.²²⁹ It terminates at the heaviest stable isotope it can generate, namely ^{209}Bi , although the culminating s-process cycle involves ^{210}Bi , ^{210}Po , ^{206}Pb , and ^{209}Pb .

Mass Loss and AGB Evolution: Stars in the AGB phase lose mass rapidly by expelling large amounts of matter including dusts of silicates or carbon, helping to enrich the ISM. The possible causes of mass loss include high stellar wind, helium shell flashes, and pulsations of the envelope in LPVs. In addition, as the stars expand, the surface gravity diminishes, lowering the binding energy.

Stars of ZAMS mass $> 8 \text{ M}_\odot$ will eventually supernova, and will be discussed later. Stars of ZAMS mass $< 8 \text{ M}_\odot$ continue to synthesize C and O in the He-burning shell. In the core, increasing density leads to rising and dominance of pressure by electron degeneracy. For stars of ZAMS mass $< 4 \text{ M}_\odot$, the C and O in the core will never burn. For stars of ZAMS mass 4 to 8 M_\odot , if there had been no mass loss, the core would reach a point where it could not maintain hydrostatic equilibrium, despite the additional electron degeneracy pressure, and a catastrophic core collapse would occur. Instead however, the high mass loss rate has reduced mass sufficiently to prevent collapse. The core synthesizes more oxygen, neon, and magnesium (becoming a ONeMg core) and stellar masses at this stage are below the Chandrasekhar limit of 1.4 M_\odot .

The intense **superwind** of around $10^{-4} \text{ M}_\odot \text{ yr}^{-1}$ may be responsible for **OH/IR sources**. OH radicals (which are unstable molecules) are detected by infrared emissions of **maser** transitions from a long-lived metastable state to ground state. (Galactic superwinds with OH megamasers have also been described.²³⁰)

Post-Asymptotic Giant Branch (post-AGB): Eventually, the cloud surrounding the supergiant star thins and a supergiant star typically of F or G spectral class becomes apparent. The track for supergiants turns blueward, moving nearly horizontally (constant luminosity). The remainder of the star's envelope is eventually expelled, the H and He burning shells are exhausted, and the star drops rapidly in luminosity until it becomes a **white dwarf star (WD)**. A white dwarf has a degenerate C-O core surrounded by a thick layer of residual H and He. These are the final remnants of ZAMS stars of $M < 8 \text{ M}_\odot$. For instance, a star of 0.6 M_\odot at ZAMS becomes a WD with radius 0.0285 R_\odot .

Planetary Nebula: A glowing ionized expanding shell of gas may be seen around a WD progenitor, termed a *planetary nebula* (a misnomer left over from before 1785). These are often multi-colored due to emissions in visible light of multiply ionized species. For example, blue-green arises from 500.68 nm and 495.89 nm forbidden lines of [O III], etc.

"Planetary nebulae play a crucial role in the chemical evolution of the galaxy, returning material to the interstellar medium that has been enriched in heavy elements and other products of nucleosynthesis (such as carbon, nitrogen, oxygen and calcium). In more distant galaxies, planetary nebulae may be the only objects that can be resolved to yield useful information about chemical abundances."²³¹

An image of a prominent planetary nebula follows.

²²⁹ ibid.

²³⁰ Baan WA, et al, "Molecular outflows in powerful OH megamasers", *Astrophysical Journal, Part 1* vol. 346, Nov. 15, 1989, p. 680-689, <http://adsabs.harvard.edu/abs/1989ApJ...346..680B>

²³¹ http://en.wikipedia.org/wiki/Planetary_nebula



The Helix Nebula, NGC 7293, a planetary nebula showing a pre-WD in the center emitting UV that excites the surrounding gas and produces inner rim cometary knots & filaments.

Distance 650 ly (200 pc), width of uncropped image 27 arcmin = 5.1 ly = 1.6 pc
(May 2003, cropped NASA HST image)²³²

Stellar Clusters

Stellar or star clusters are groups of stars that are not galaxies. They are often the building blocks of galaxies. They are very useful in testing theories of stellar evolution. It is generally safe to assume that all the stars in a cluster formed at about the same time from similar starting compositions and with the same distance modulus. The distance modulus can be determined by matching the main sequence of the cluster to a main sequence calibrated in absolute magnitude, a process called by the misnomer **spectroscopic parallax**, also referred to as **main-sequence fitting** (discussed further below).

Metallicity Z, [Fe/H], and [O/H]

In astronomy, the term “**metals**” is used (somewhat perversely!) for atoms and nuclei other than H and He. The term **metallicity** is given by **Z**, where Z=mass fraction of nuclei having atomic number > 2. The Sun has Z \approx 0.016 to 0.018. Recall that X is the mass fraction of H (H kg/total kg), and Y is the mass fraction for He (He kg/total kg), and that the mass fractions X + Y + Z = 1.

²³² <http://hubblesite.org/newscenter/archive/releases/2003/11/image/a/>
Page 94 of 249

Another measure of metallicity, expressed relative to the Sun, is “[Fe/H]”, defined by:

$$[\text{Fe}/\text{H}] \equiv \log_{10} \left[\frac{\left(\frac{N_{\text{Fe}}}{N_{\text{H}}} \right)_{\text{star}}}{\left(\frac{N_{\text{Fe}}}{N_{\text{H}}} \right)_{\odot}} \right] \equiv \log_{10} \left(\frac{N_{\text{Fe}}}{N_{\text{H}}} \right)_{\text{star}} - \log_{10} \left(\frac{N_{\text{Fe}}}{N_{\text{H}}} \right)_{\odot}$$

Here N are number densities (atoms per unit volume). Similar definitions apply to [O/H].

The metallicity [Fe/H] of the Sun is 0.0 as hereby defined. Here,²³³

$$\log_{10} \left(\frac{N_{\text{Fe}}}{N_{\text{H}}} \right)_{\odot} \cong -4.33$$

Stellar Populations

Stars are grouped on the basis of **metallicity** (given by Z or [Fe/H]. The hypothetical earliest stars, produced soon after the big bang, consisted when first formed almost entirely of H and He. They therefore began with no “metals” content, so Z = 0. These are the hypothesized metal-poor **Population III** stars. More recently formed but still very old and metal-poor stars with Z≈0 are termed **Population II**. Newer stars, which had more metals available to incorporate, have mass fraction values as high as Z≈0.03 and are termed metal-rich **Population I** stars. The Sun is in this population I.

Definitions of Open and Globular Stellar Clusters

There are two types of star clusters: **globular clusters**, which are tight groups of hundreds of thousands of very old stars that are gravitationally bound, and **open clusters**, which are more loosely clustered groups of stars, generally contain fewer than a few hundred members, and are often very young.²³⁴

Open clusters, also confusingly called **galactic clusters** (these are not clusters of galaxies), tend to be smaller and contain younger Pop. I stars, numbering a few thousand. They have been found in spiral and irregular galaxies. Many are inherently unstable... There are over 1000 in the Milky Way. “Open clusters generally survive for a few hundred million years. In contrast, the more massive globular clusters of stars exert a stronger gravitational attraction on their members, and can survive for many billions of years.”²³⁵ “Open clusters become disrupted over time by the gravitational influence of giant molecular clouds as they move through the galaxy, but cluster members will continue to move in broadly the same direction through space even though they are no longer gravitationally bound; they are then known as a **stellar association**, sometimes also referred to as a **moving group**.²³⁶

Globular Clusters GC

A globular cluster is a spherical collection of stars that orbits a galactic core as a satellite.²³⁷ It is not a galaxy but is tightly bound by gravity, giving spherical shapes and high stellar densities. Examples include Messier 80. They are found within galaxies in the galactic halo. They consist of tens to hundreds of thousands of stars. The M13 cluster like most globular clusters consists of very old Pop. II stars. .

Large collections of cluster stars are often studied in **color-magnitude diagrams**, plotting for instance V [visual band magnitude] on the ordinate and B - V on the abscissa. The patterns of stars trace out various of the stages of stellar evolution:

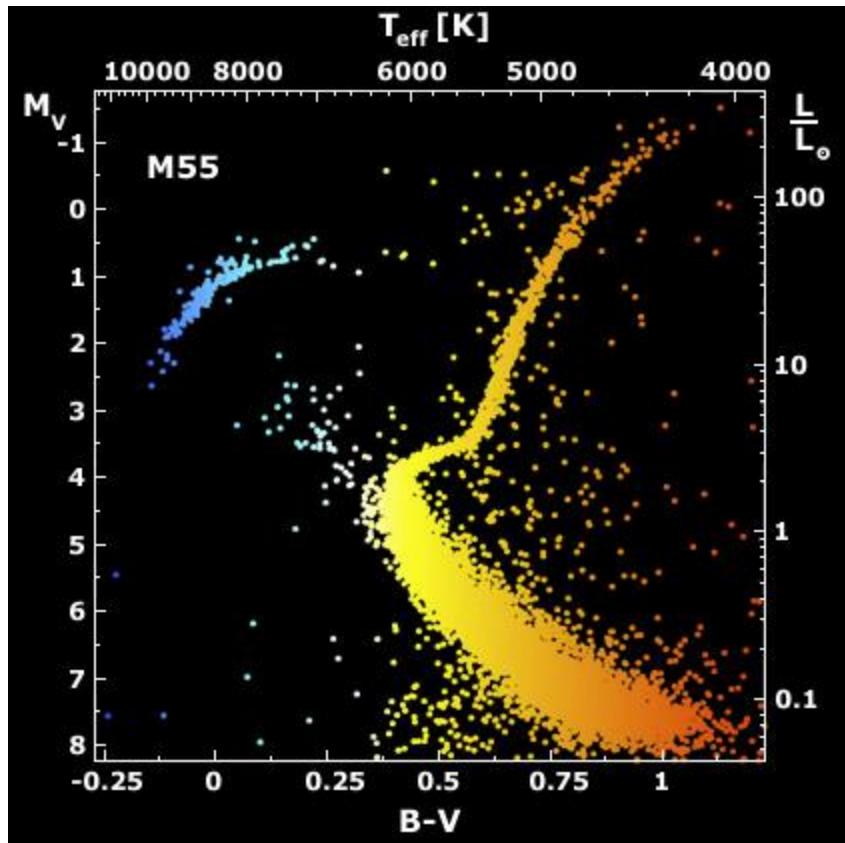
²³³ <http://burro.astr.cwru.edu/Academics/Astr222/Galaxy/Structure/metals.html>

²³⁴ http://en.wikipedia.org/wiki/Star_cluster

²³⁵ http://en.wikipedia.org/wiki/Open_cluster

²³⁶ http://en.wikipedia.org/wiki/Star_cluster

²³⁷ http://en.wikipedia.org/wiki/Globular_cluster



Color Magnitude Diagram for Globular Star Cluster M55, NASA, February 2001.

The Main Sequence, Blue Stragglers, Turnoff to the SGB, SGB, RGB, and HB including Blue Giants are apparent. The E-AGB and Post-AGB are apparently not included. (NASA)²³⁸

The position of the sharp upper turn-off from the main sequence toward the subgiant and red giant branches is a useful index of the cluster's age. The more time that has elapsed, the lower will be the average masses of stars remaining on the main sequence, and thus the turnoff will be toward the "later" spectral types. A series of turnoff curves vs. estimated age is shown on *IMA2* p. 478. A few **blue stragglers** may be present above and to the left of the turnoff.

The relative number of stars at various positions in the MS versus in other evolutionary stages at a given time (a pattern termed an **isochrone**) also gives indication of the cluster age.

Stars in the **Hertzsprung Gap** are moving rapidly through this often sparsely populated SGB location toward the RGB stage. For similar reasons, namely the brevity of these stages, few stars may be apparent in the AGB and Post-AGB phases as stars evolve toward the WD stage.

²³⁸ <http://apod.nasa.gov/apod/ap010223.html>; see also *IMA2* p. 476
Page 96 of 249

Stellar Pulsation (Chapter 14)

This chapter deals primarily with pulsating variable stars (i.e., those that intrinsically and repetitively pulsate over significant amplitudes). I can only scratch the surface of this rich and fascinating subject matter, a professional focus for Prof. Szkody.

This chapter excludes:

- *eruptive variables*—stars that undergo cataclysmic changes (these arise from a WD primary with a mass transferring secondary, and include novae, dwarf novae, novalike eruptions, polars and low accretion rate polars LARPs, and AM CVn stars)²³⁹ or catastrophic changes (such as supernovae); stars that emit eruptive flares, mass ejections (such as Wolf-Rayet stars), or X-ray bursts or gamma-ray bursts, and
- *extrinsic variable stars*—stars whose variability results largely from their rotation (by having prominent starspots, polar brightness variation, ellipsoidal shapes, etc.), as well as eclipsing binary systems).²⁴⁰

The **International Variable Stars Index** (AAVSO VSX) is a useful place to find virtually all major variable stars listed by type and searchable by applying various criteria, though tricky to use.²⁴¹

Key Observations of Pulsating Stars

David Fabricius first noted in 1595 the variation in the brightness of o Ceti (**omicron Ceti**), which he termed **Mira**, meaning wonderful. (Earlier observations of this star's variability are uncertain.) Its period of 11 months was found by Johannes Holwarda, making this a **long-period variable** (those having periods of 100 to 700 days), and it is now the prototype of the *Mira Cet variable type*.²⁴² Mira is actually a binary star system at ~90 pc, and Mira A is the much larger and variable component. Mira A is a red giant and said to be an Asymptotic Giant Branch (AGB) star, in the thermally pulsing (TP-AGB) phase.²⁴³ The oscillation period averaging 333 days varies from about 328 days to about 336 days over a longer period of about 40 years.²⁴⁴ The variation within a single period extends a wide range, from magnitude ~2 to ~10. The variation is due to the star expanding and contracting, with changes in brightness resulting.

Another pulsating star, whose variability was discovered in 1784, is **δ Cephei**—the prototypical **classical Cepheid**. It has a period of 5d 8hr 48min and magnitude variation of about 0.7 magnitudes. Henrietta Swan Leavitt (1868–1921) working at Harvard, identified more than 2400 classical Cepheids with periods from 1 to 50 days, mostly in the SMC. She plotted and observed the close relationship between period and luminosity. A more recent 1968 graph of this **Period-Luminosity (PL) relation** follows:

This graph at right shows the Period-Luminosity relationship for classical

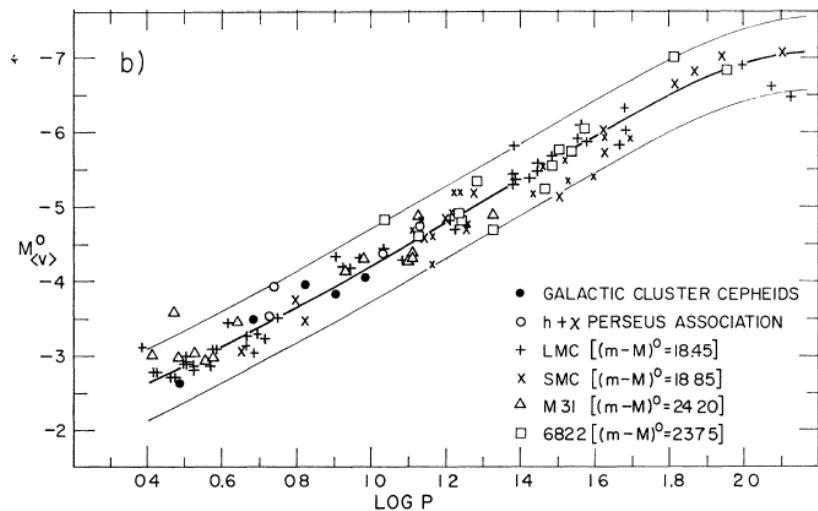


FIG. 1.—The composite period-luminosity relation at mean intensity in *B* and *V* wavelengths derived from the sources indicated at the lower right. The absolute calibration was made by using the nine Cepheids of the galactic system shown as open and filled circles. The photographic data from the SMC are plotted with smaller crosses than the Gascoigne and Kron photoelectric data.

²³⁹ <http://heasarc.gsfc.nasa.gov/docs/objects/cvs/cvstext.html>

²⁴⁰ International Variable Stars Index:

- <http://www.spaceacademy.net.au/library/guides/varstar/varstar.htm>
- http://en.wikipedia.org/wiki/Variable_star

²⁴¹ <http://www.aavso.org/vsx/index.php?view=search.top>

²⁴² http://en.wikipedia.org/wiki/Mira_variable

²⁴³ <http://en.wikipedia.org/wiki/Mira>

²⁴⁴ <http://arxiv.org/abs/0902.3604>

Cepheid variables in MW clusters Cepheids, in the open cluster h and χ Persei, and in the SMC, LMC, M31, and NGC 6822 (Barnard's Galaxy). (Sandage and Tamman 1968²⁴⁵, graph for M_B omitted)

According to *IMA2* p. 486, this data is fitted to the equation

$$M_{(V)} = -2.81 \log_{10} P_d - 1.43,$$

where P_d is the period in days and $M_{(V)}$ is the average absolute V magnitude. Sandage and Tamman state, "The data are consistent with the view that the P-L relations are the same from galaxy to galaxy in both slope and zero point." Tighter results are reported using the infrared H band where extinction is lower.

An even tighter correlation arises from incorporation of an infrared color index (J band – K_s band), so that the best relation comparing H band absolute magnitude, period P in days, and the average color index $(J - K_s)$ is reported by *IMA2* (p. 487) to be

$$H = -3.428 \log_{10} P + 1.54(J - K_s) + 15.637$$

This is called a **Period-Luminosity-Color (PLC) relation**.

Persson et al (2004) give recent a PLC calibration of Cepheids in the LMC using near-infrared light curves as follows:²⁴⁶

The graph to the right is a PLC plot for Cepheids in the Large Magellanic Cloud. Here, the vertical axis *mag* incorporates the $W(K_s)$ Wesenheit parameter and (according to *IMA2* p. 488) is of the form $H - c(J-K_s)$. These corrections provide the tightest correlation with Log Period. (Persson SE et al 2004)

(I have not yet taken time to study the complex relationships described in these technical articles.)

Classical Cepheids, which are supergiant stars of high luminosity class Ib (i.e., a less luminous supergiant, thus having luminosity $\sim 1000 L_\odot$), provide one of the standard candles needed to estimate intergalactic distances.

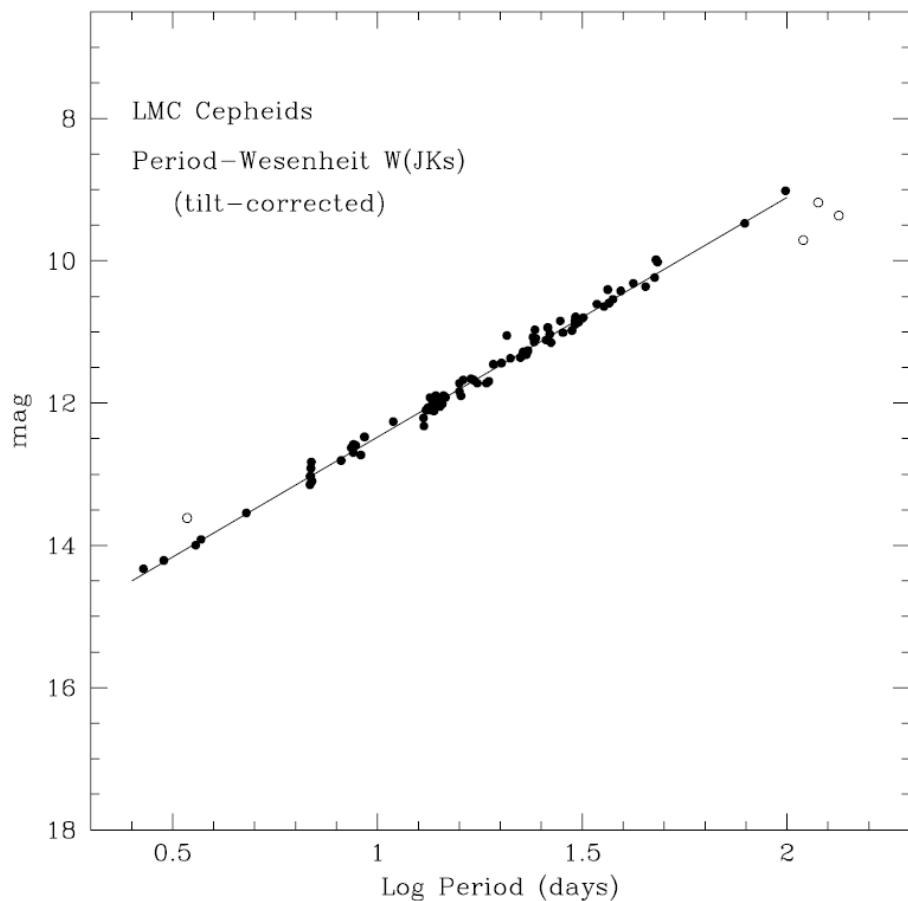


FIG. 5.—The $W(K_s)$ PW parameter plotted against $\log P$. The Wesenheit parameters were corrected for the average tilt of the distribution in analogy with Fig. 4; the coefficients are given in Table 7.

²⁴⁵ Sandage AR & Tamman GA, "A Composite Period-Luminosity Relation for Cepheids at Mean and Maximum Light", *Astrophysical Journal*, vol. 151, p.531, 1968.

²⁴⁶ Persson SE, et al, "New Cepheid Period-Luminosity Relations for the Large Magellanic Cloud: 92 Near-Infrared Light Curves". *The Astronomical Journal* 128 2239, 2004,

<http://iopscience.iop.org/1538-3881/128/5/2239>

The Pulsation Hypothesis for Brightness Variations

Harlow Shapley (1885–1972) showed that the variability of classical Cepheid stars is not attributable to their being in binary systems, and suggested that the variability of brightness and T arises from **radial pulsation** of a single star. Arthur Eddington (1882–1944) provided the theoretical underpinning of this theory. For δ Cephei, the full excursion or change of brightness (which might be expressed as ΔL or ΔM_{bol}) is due primarily to the ~ 1000 K excursion in surface temperature (ΔT_{eff}), with a lesser contribution to ΔL or ΔM_{bol} due to excursion in radial size ΔR (increased r of course increases radiating surface). $\Delta R/R_{\text{mean}}$ is only about 5 to 10%, though ΔR may be greater than $2R_{\odot}$ in these giant stars. The spectral type varies during the pulsation cycle from F5 at the peak to G2 at the nadir.

According to Martin Schwarzschild's 1938 observed and derived data for δ Cephei (shown in graphs below),²⁴⁷ the T_{eff} and M_{bol} peaks are at almost simultaneous phases as expected (phases 0.98 and 1.00, respectively) and occur somewhat past when the star is most rapidly expanding (when \dot{r} is maximally positive, at phase 0.90) and when the star passes through the minimum radius (at phase 0.92).

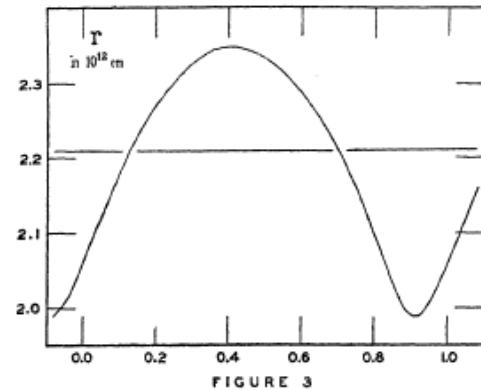


FIGURE 3

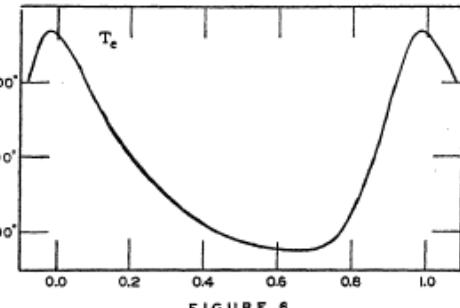


FIGURE 6

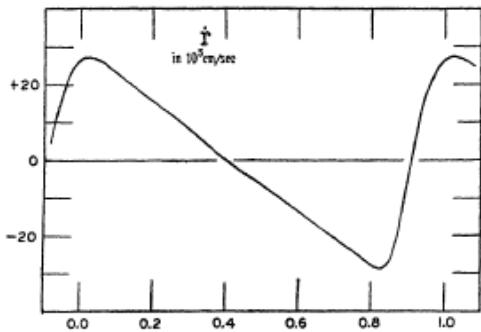


FIGURE 4

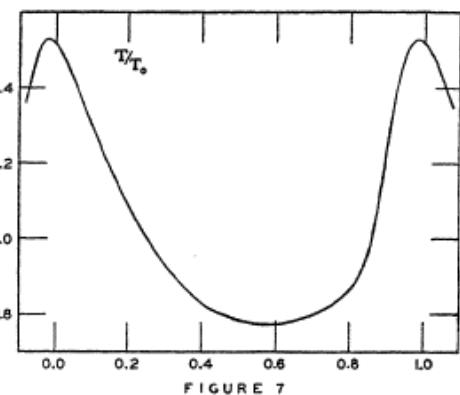


FIGURE 7

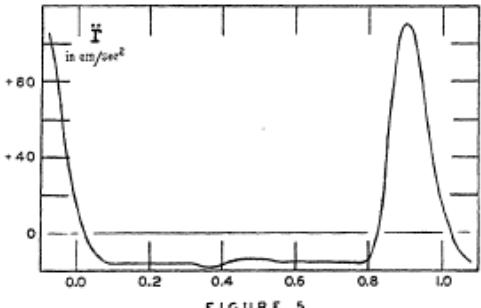


FIGURE 5

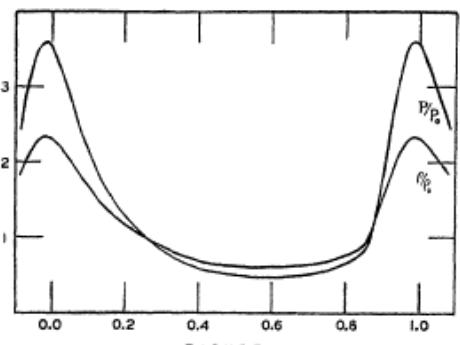


FIGURE 8

Variability in δ Cephei over one cycle:

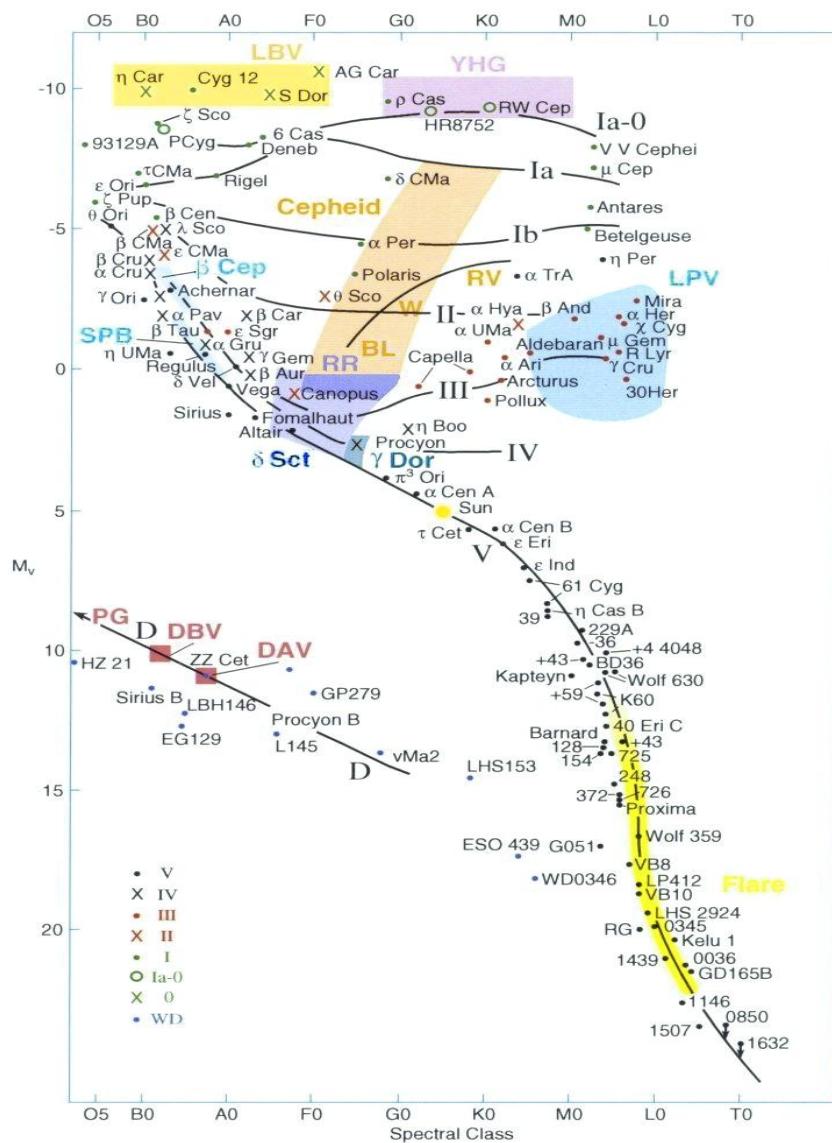
L-R, top to bottom:

Stellar Radius r , Surface Temp T_{eff} , Surface Velocity \dot{r} ,
 T/T_{\star} , Surface Acceleration \ddot{r} , and Ratios of P/P_{\star} and ρ/ρ_{\star} .²⁴⁸

²⁴⁷ Martin Schwarzschild, "On the Light Curve of δ Cephei" *Harvard College Observatory Circular*, vol. 431, pp.1-13, 12/1938 For exact phases, see Table VI. [Martin Schwarzschild (1912 - 1997) is son of Karl S.]

²⁴⁸ ibid.

The Instability Strip and H-R Regions Where Pulsating/Variable Stars Are Found



H-R Diagram showing Regions of Stellar Variability or Pulsatility.²⁴⁹
 (Note that spectral classes depicted here are aligned vertically,
 not always the case for H-R diagrams)

Regions designated here include the following (not all of which exhibit regular pulsation)

- LBV=Luminous Blue Variables (S Doradus);**
- YHG=Yellow Hypergiant variables;**
- Classical Cepheid (Type/Population I Cepheids, including δ Cephei);**
- LPV=Long Period Variables (Mira variables);**
- BL, W, & RV Cepheids (Type/Population II Cepheids, metal poor, incl. BL Herculis, W Virginis, and RV Tauri stars)**
- RR=RR Lyrae stars;**
- δ Sct= Delta Scuti stars;**
- γ Dor=Gamma Doradus stars;**
- β Cep=Beta Cephei stars;**
- SPB= Slowly Pulsating B stars (53 Persei stars);**
- DAV=White dwarf variables (aka ZZ Ceti);**
- DBV=White dwarf variables (aka V777 Her);**
- PG=PG 1159 stars (very hot white dwarf variables);**
- Flare=Flare stars (mostly class M red dwarfs).**

²⁴⁹ image (modified by MCM) fr. <http://stars.astro.illinois.edu/sow/hrd.html>, fr. JB Kaler, *The Cambridge Encyclopedia of Stars*, Cambridge U Press, Nov 2006

Obviously, the detailed classification of variable stars is quite complex. The types of variable stars and full descriptions are listed here²⁵⁰. This summary mentions other variable star types not listed in this table.

Here are some of the main types of pulsating variables (plus a few ringers), listed in top-down order as shown on the HR Diagram above (definitions and quoted material from AAVSO website (www.aavso.org) unless otherwise noted):

Type of Variable Star	Period II (some stars are multiperiod or nonperiodic)	Notes (some edited with added abbreviations)
LBV=Luminous Blue Variables (aka S Doradus variables)	not regular pulsations	“... are very bright, blue, hypergiant variable stars named after S Doradus, the brightest star of the LMC. They exhibit long, slow changes in brightness, punctuated by occasional outbursts in brightness during substantial mass loss events (e.g. Eta Carinae, P Cygni). They are extraordinarily rare. The General Catalogue of Variable Stars [GCVS] only lists 20 objects as SDor.” ²⁵¹
YHG=Yellow Hypergiant variables	? not regular pulsations	“...form a class of very luminous yellow supergiants that show strongly turbulent photospheres and high rates of mass loss... With effective temperatures of only 4000–7000 K, they occupy the cooler half of the Hertzsprung gap, which is the wide yellow band that separates red supergiants from luminous blue variables (LBVs or S Doradus variables)... It is now believed that yellow hypergiants are the cores of massive stars that have recently evolved out of a much cooler red supergiant state...” ²⁵²
Classical Cepheid: Population I variables (including δ Cephei)	1 to 70 days <i>IMA2</i> : 1-50 days	“Light variations from 0.1 to 2 magnitudes. These massive stars have high luminosity and are of F spectral class at maximum, and G to K at minimum. The later the spectral class of a Cepheid, the longer is its period. Cepheids obey the period-luminosity relationship.” [AAVSO] Radial mode.
LPV=Long Period Variables (aka Mira Variables)	30-1000 days <i>IMA2</i> : 100-700 days	“pulsating red giants or supergiants with periods ranging from 30-1000 days. They are usually of spectral type M, R, C or N. There are two subclasses; Mira and Semiregular: Mira—these periodic red giant variables vary with periods ranging from 80 to 1000 days and visual light variations of more than 2.5 magnitudes. Semiregular—These are giants and supergiants showing appreciable periodicity accompanied by intervals of semiregular or irregular light variation. Their periods range from 30 to 1000 days, generally with amplitude variations of less than 2.5 magnitudes.” [AAVSO] Radial mode.

²⁵⁰ Classification of variable stars:

- <http://www.aavso.org/types-variables>
- http://en.wikipedia.org/wiki/Variable_stars

²⁵¹ http://en.wikipedia.org/wiki/Luminous_blue_variable_star

²⁵² <http://iopscience.iop.org/0004-637X/560/2/934/fulltext/53874.text.html>

BL, W, & RV Cepheid Variables: Type/Population II Cepheid variables, metal poor, incl. BL Herculis, W Virginis, and RV Tauri	1 to 150 days BL Her: 1 to 5 days W Virginis: 10 to 20 days RV Tauri: 30 to 150 days ²⁵³	<ul style="list-style-type: none"> • RV Tauri: “These are yellow supergiants having a characteristic light variation with alternating deep and shallow minima. Their periods, defined as the interval between two deep minima, range from 30 to 150 days. The light variation may be as much as 3 magnitudes. Some of these stars show long-term cyclic variations from hundreds to thousands of days. Generally, the spectral class ranges from G to K.” • BL Her or Herc: “[of Population II Cepheids,] the short-period stars are called BL Herculis stars, the long-period stars generally have RV Tauri characteristics—alternating deep and shallow minima. • W Virginis: “That leaves the medium-period stars, which are called W Virginis stars.” <p>[all AAVSO] Radial mode (at least for W Virginis)</p>
RR=RR Lyrae variables	0.05 to 1.2 days <i>IMA2:</i> 1.5-24hr	<p>“These are short-period (.05 to 1.2 days), pulsating, white giant stars, usually of spectral class A. They are older and less massive and less luminous than Cepheids. The amplitude of variation of RR Lyrae stars is generally from 0.3 to 2 magnitudes.” [AAVSO] Radial mode</p>
δ Sct= Delta Scuti variables	0.03 to 0.3 day <i>IMA2:</i> 1-3 hr	<p>“...reside near the point where the instability strip crosses the main sequence in the HR diagram. Stars in a variety of evolutionary states—including pre-main sequence stars—can lie within the instability strip, so long as they have spectral types between (roughly) F8 and A2, and luminosity classes between V (dwarf) and III (subgiant). For stars with solar metal abundances, this corresponds to masses between about 1.5 and 2.5 solar masses, and between 1.0 and 2.0 solar masses for metal-poor stars. They are all short period stars, with individual periods lying in the range of 0.03 to 0.3 day... A few delta Scuti stars (like FG Virginis and XX Pyxidis) are known to have a dozen or more of these non-radial modes excited at once, and can be detected via the analysis of very precise time-series photometry.” [AAVSO] Radial & nonradial modes</p>
γ Dor=Gamma Doradus variables	0.4 to 3 days	<p>“...class of pulsating stars, the prototype of which is the bright, early, F-type dwarf γ Doradus. These stars typically have between 1 and 5 periods ranging from 0.4 to 3 days with photometric amplitudes up to 0.1 mag in Johnson V [band]. The mechanism for these observed variations is high-order, low-degree, nonradial, gravity-mode pulsation.”²⁵⁴</p> <p>Nonradial modes</p>

²⁵³ Periods for types of Pop II Cepheids from <http://www.jstor.org/stable/10.1086/341698>

²⁵⁴ <http://www.jstor.org/stable/10.1086/316399>

β Cep=Beta Cephei variables	0.1 to 0.3 days <i>IMA2:</i> 3 to 7 hr	"Beta Cephei variables are sometimes known as β Canis Majoris stars... But, in the General Catalogue of Variable Stars, they are classified as BCEP ... They have spectral types of B0-B2 III-V, corresponding to temperatures of 20-30,000 K, i.e. they are near the top of the main sequence on the Hertzsprung-Russell diagram. Their masses are typically 10-20 times that of the sun. They have periods of 0.1 to 0.3 days and, though their radial velocity amplitudes are quite large, their visual light amplitudes are very small, typically only a few hundredths of a magnitude... Satellite observations show that, in ultraviolet light, where most of the flux from these hot stars is found, their amplitudes are up to a magnitude... But the small visual amplitude explains why the measurement of radial velocity variations was the observational method of choice. BW Vulpeculae, the largest-amplitude β Cephei star, has a vigorous pulsation amplitude of 200 km/s..." [AAVSO] Radial & nonradial modes
SPB= Slowly Pulsating B stars (aka 53 Persei variables)	0.5-5 days ²⁵⁵	"ζ [Zeta] Cas is a member of an unusual group of variable stars known as "Slowing Pulsating B" (SPB) stars. It shows a pulsation frequency of 0.64 per day (or once every 1.56 days) and displays a weak magnetic field with a strength of roughly 3.35×10^{-2} T, which varies with a period of 5.37 days. This likely matches the rotation rate of the star, which, when combined with the low projected rotational velocity, indicates the star may be seen nearly pole-on." ²⁵⁶ "There are a large number of β Cephei stars among the brightest stars in the sky.... Starting about [1980], there was much interest in the newly-discovered slowly-pulsating B stars, the 53 Persei stars, and pulsating Be stars..." Non-radial modes [AAVSO] ²⁵⁷
DAV=White dwarf variables (aka ZZ Ceti variables)	<i>IMA2 p.</i> 561: 100 to 1000 sec	"luminosity varies due to non-radial gravity wave [g-mode] pulsations" "The DAV stars (commonly called ZZ Ceti stars) have temperatures around 10 000 – 12 000 K," ²⁵⁸ D = degenerate, V = variable Nonradial g-modes
DBV=White dwarf variables (aka V777 Her)	minutes ²⁵⁹	"g-mode oscillations driven by the helium partial ionization zone" [<i>IMA2 p.</i> 562] "The DBV stars have temperatures in the range ~22,000 – 27,000 K" ²⁶⁰ D = degenerate, V = variable Presumably nonradial modes
PG=PG 1159 variables (aka GW Vir, very hot white dwarf variables). Divided into DOV and PNNV stars	300 to 3,000 sec	pre-degenerate, hydrogen-deficient atmosphere, surface temperatures between 75,000 K and 200,000 K... These stars vary slightly (5–10%) in brightness due to non-radial gravity wave pulsations within themselves. They vibrate in a number of modes simultaneously, with typical periods between 300 and 3,000 seconds." ²⁶¹ nonradial modes ²⁶²

²⁵⁵ <http://www.ster.kuleuven.ac.be/~peter/Bstars/>

²⁵⁶ http://en.wikipedia.org/wiki/Zeta_Cassiopeiae

²⁵⁷ http://www.aavso.org/vsots_betacep

²⁵⁸ www.star.ac.za/course-resources/local/david-buckley/var2.pdf

²⁵⁹ Period of GD 358 given by <http://adsabs.harvard.edu/abs/1982ApJ...262L..11W>

²⁶⁰ ibid.

²⁶¹ http://en.wikipedia.org/wiki/PG_1159_star

Flare=Flare stars (mostly class M red dwarfs including UV Ceti)	not regular pulsations, flares last seconds to minutes	"UV Ceti, however, is an extreme example of a flare star that can boost its brightness by five times in less than a minute, then fall somewhat slower back down to normal luminosity within two or three minutes before flaring suddenly again after several hours. In 1952, UV Ceti was observed flaring to 75 times its normal brightness in only 20 seconds." ²⁶³ "...a variable star that can undergo unpredictable dramatic increases in brightness for a few minutes. It is believed that the flares on flare stars are analogous to solar flares in that they are due to magnetic reconnection in the atmospheres of the stars. The brightness increase is across the spectrum, from X rays to radio waves... The best-known flare star is UV Ceti, discovered in 1948. Today similar flare stars are classified as UV Ceti type variable stars (using the abbreviation UV) in variable star catalogs such as the General Catalogue of Variable Stars... Most flare stars are dim red dwarfs, although recent research indicates that less massive brown dwarfs might also be capable of flaring. The more massive RS Canum Venaticorum variables (RS CVn) are also known to flare, but it is understood that these flares are induced by a companion star in a binary system which causes the magnetic field to become tangled." ²⁶⁴
--	--	---

The MW contains several million pulsating stars out of several hundred billion, so variability may be relatively transient. The stars with substantial pulsation are not generally located on the Main Sequence, but are found in the **Instability Strip** (shown above extending from classical Cepheids to RR Lyrae and δ Sct stars) and in other regions noted. As star evolve through the instability strip, they begin to pulsate, and stop when they leave it.

Ultimately, most if not all stars including the Sun²⁶⁵ exhibit some degree of variability in T and luminosity, etc., but the fluctuations tend to be relatively small and inconspicuous compared to the stars designated as *variable*.²⁶⁶

The Physics of Stellar Pulsation

Period-Density Relation

Radial stellar oscillations are due to sound wave resonances in the interior (*IMA2* p. 491). The period of pulsation Π (symbol Pi is used to avoid confusion with pressure P). The time for sound to cross the diameter of a star is crudely estimated using adiabatic sound speed Vs at *IMA2* p. 492, yielding a rough estimate of the pulsation period for the Fundamental radial mode.

$$\Pi \approx \sqrt{\frac{3\pi C_v}{2G\rho C_p}}$$

where ρ = constant average density (unrealistic) and C_v = Specific Heat at constant volume and C_p = Specific Heat at constant pressure. Therefore, period is approximately inversely proportional to the square root of mean density. For a typical Cepheid, the period this formula estimates is 10 days, which is in the right

²⁶² <http://adsabs.harvard.edu/abs/2004astro.ph..9243N>

²⁶³ <http://www.solstation.com/stars/luy726-8.htm>

²⁶⁴ http://en.wikipedia.org/wiki/Flare_star

²⁶⁵ http://science.nasa.gov/science-news/science-at-nasa/2010/05feb_sdo/

²⁶⁶ <http://cseigman.com/text/stars/variables.htm>

ballpark. This relationship helps to explain why periods decrease as one moves down the instability strip (toward lower luminosity) from tenuous supergiants to dense white dwarfs.

Radial vs. NonRadial Modes of Spherical Oscillations (Spherical Harmonics)

Radial Modes: The radial mode stellar pulsations are essentially standing sound waves, analogous to those of an open-ended organ pipe. “The pressure and velocity perturbations in a standing acoustic wave are 90° out of phase. Thus at the closed end of the pipe, the velocity vanishes and the pressure fluctuations are maximal. Similarly, at the center of a radially pulsating star, the radial velocity has a node and the pressure has an antinode. At the open end of the organ pipe and at the surface of a star, the situation is reversed: the pressure fluctuations vanish and those of the velocity are maximal...”²⁶⁷ The standing wave for each radial mode has a velocity node at the star’s center and a velocity antinode at the surface. In the **fundamental mode** (*first harmonic*), the gases throughout the star are moving outward or inward with radial symmetry and there is no other node than at the center. For this mode, the maximal fractional displacement of the gas from its equilibrium position rises steadily as one moves from the velocity node at the center to the surface.

For the *first overtone* (second harmonic), there is a spherical velocity node at $\sim 0.6R_*$, so that gas moves inward inside this node while gas outside it moves outward, and conversely (all with radial symmetry). The second overtone has 2 concentric spherical nodes, the third overtone has 3 concentric spherical nodes, etc.

For radial modes, most of the motion takes place at or near the surface, with lesser movement deep inside the star. For example, the maximal fractional displacement from equilibrium for the fundamental mode at $0.5 R_*$ is only 7% of that at the surface ($1 R_*$). For higher overtones, the disproportion between surface versus deep displacement is even greater. The vast majority of classical Cepheids and W Virginis stars oscillate in fundamental radial mode. RR Lyrae and possibly LPVs like Mira stars oscillate in fundamental or first overtone modes, sometimes combined (*IMA2* p. 494).

“Radial modes often produce the largest changes in a star’s radius, leading to large changes in luminosity, temperature, and radial velocity... The simplest radial mode is the fundamental mode, in which the entire atmosphere of the star expands/contracts in unison. **Nonradial modes** are more complicated, no longer radially symmetrical, as some parts of the star’s surface move inwards while other parts move outwards... Examples of nonradial pulsators include pulsating white dwarfs, Gamma Doradus stars, Beta Cepheids, and Delta Scuti stars.”²⁶⁸ (For further discussion on nonradial modes, see later.)

Mechanism Powering Stellar Standing Waves

Arthur Eddington proposed that pulsating stars are thermodynamic heat engines. This is illustrated by means of a P-V thermodynamic diagram, which shows a closed loop as one goes through a complete period. If the integral about the complete path $\oint P dV > 0$ (the path is clockwise), the gases of the star do net $P dV$ work through the cycle (in a **driving layer**), whereas if the integral is < 0 (the path is counterclockwise), the gases are acting as a **damping layer**. Oscillations will grow if the total work (sum of integrals for all the layers) > 0 . At an equilibrium value of pulsation amplitude, the total work done by all layers is zero. The driving layer must absorb heat at around the time of maximal compression... The core of the star usually exhibits too little pulsation amplitude to power the pulsations. (However, oscillations in a star’s nuclear energy generation rate contribute to the Eddington limit on main sequence stellar mass of about $90 M_\odot$. Eddington suggested that a layer inside the star becomes more opaque in compression, restricting the flow of energy to the surface—with expansion and reduced compression, the opacity diminishes and energy flow to the surface resumes at the higher rate. This layer is referred to as **Eddington’s Valve**. Only in special circumstances does opacity increase with compression, explaining the relative rarity ($1:10^5$) of pulsating stars. This region was identified as a **partial ionization zone**. Here, work done by the gases causes increased ionization, increasing Kramers’ opacity κ . The opacity decreases with decreasing density. This is the **κ -mechanism**.

An additional **γ -mechanism** reinforces the effect. Specifically, the κ -mechanism is reinforced by the γ -mechanism in a partial ionization zone. The temperature of the partial ionization zone rises less during compression than for the surrounding layers and the resulting stronger temperature gradient causes relatively increased heat flow into the cooler partial ionization zone, causing further ionization. (*IMA2* p. 497)

²⁶⁷ <http://www.astro.princeton.edu/~gk/A403/pulse.pdf>

²⁶⁸ http://www.chara.gsu.edu/~wiita/pulsating_stars2.pdf

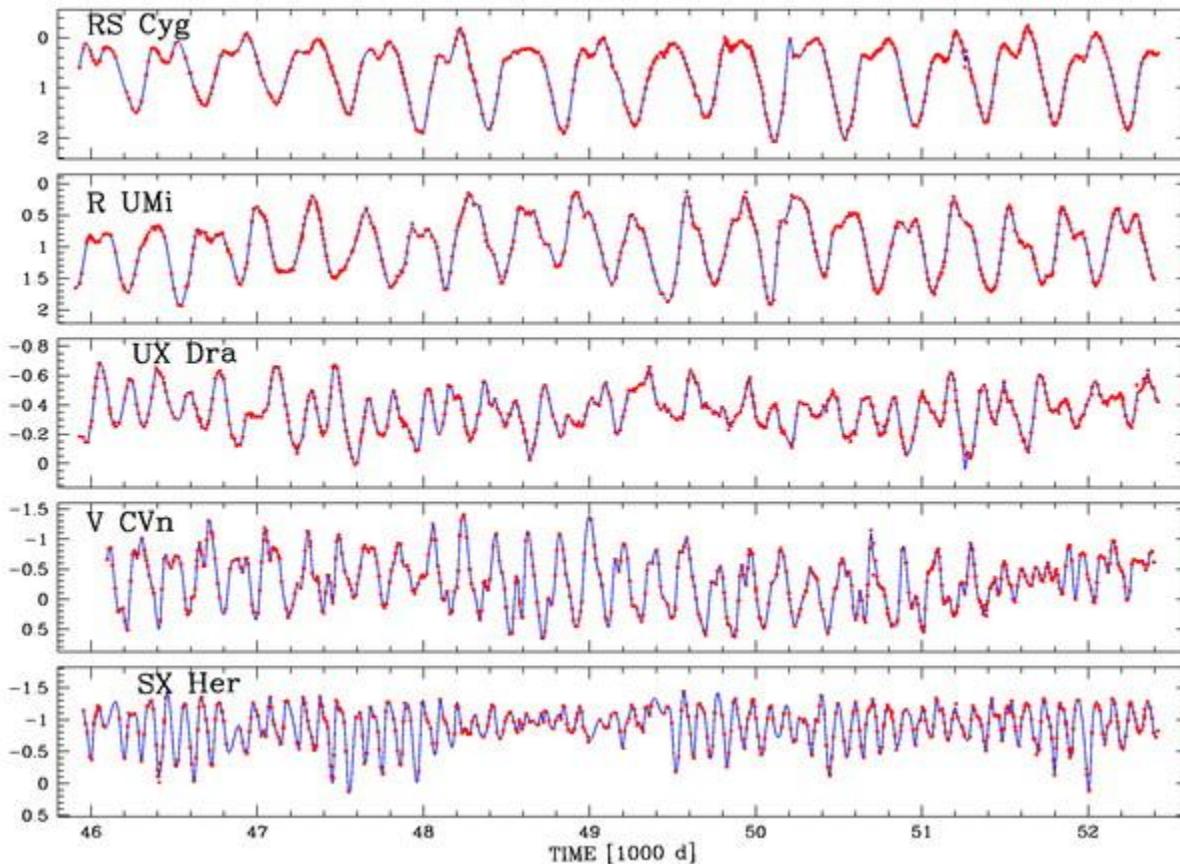
These partial ionization zones modulate the flow of energy and directly cause the pulsation of stars (*IMA2* p. 497). In most stars there are two partial ionization zones: the hydrogen partial ionization zone ($H\ I \rightarrow H\ II$ plus some $He\ I \rightarrow He\ II$) at about 1 to 1.5×10^4 K, and the helium partial ionization zone ($He\ II \rightarrow He\ III$) at about 4×10^4 K.

The location of these partial ionization zones determines the star's pulsation. At hotter temperatures of 7500 K, past the **hot blue edge** of the instability strip (on the left), the zones are too close to the surface to be effective in causing pulsation. At about 6500 K, pulsation is more effective and the first overtone may be excited. In a cooler star of 5500 K, the deeper fundamental mode may be excited. However, with still cooler surface temperatures, the **cool red edge** of the instability strip (on the right) is reached, causing damping by convection (*IMA2* p. 498).

The very hot Beta Cephei stars and early SPB Slowly Pulsating B stars are too hot for pulsation to be explained by H or He partial ionization. Instead, pulsations are attributable to an iron "bump" in opacities from iron ionization near 10^5 K (*IMA2* p. 499).

Modeling Stellar Pulsation

This is a complex subject requiring large computer resources, but has been quite successful for Cepheids. One approach utilizes nonlinear hydrodynamic models, as the star is not in hydrostatic equilibrium. When conditions are optimal in the model, the partial ionization zones will drive the oscillation. In some cases the model predicts chaotic behavior, as has been observed for some stars:



Light curves for 5 semiregular variable stars exhibiting varying degrees of chaotic variability.
Red points are actual data, curved lines are fitted to the data and smoothed.²⁶⁹

²⁶⁹ Chaos in Stellar Pulsations:

• Buchler JR, et al, "Evidence for Low-dimensional Chaos in Semiregular Variable Stars", *The Astrophysical Journal*, Volume 613, Issue 1, pp. 532-547, 09/2004. <http://iopscience.iop.org/0004->

Another approach attempts to “linearize” the differential equations. This is computationally less intensive but substantially more limited. A formula for calculating the period is illustrated (*IMA2* p. 500).

For the more sophisticated but more costly approach needed for modeling some variables, the model must be non-adiabatic, to allow heat to flow into and out of the layers, as well as non-linear. In the case where $\gamma=C_P/C_V < 4/3$, the star collapses. (I have not yet determined how valid this is and how it compares to the Chandrasekhar limit or the Schönberg-Chandrasekhar limit, etc.)

Nonradial Stellar Pulsation

I have not mastered this complex subject, and have omitted many details. Recall that with radial symmetry, parameters have the same values for a given radius r throughout the star, for all colatitudinal coordinate = inclination = polar coordinate angle θ , and for all longitudinal coordinate = azimuthal coordinate angle φ . Of course, there can be complex radial pulsation modes having spherical nodes and anti-nodes within the star, but they are all radially symmetrical.

As previously mentioned, however, spherical vibration may not be radially symmetrical.

In Laplace spherical harmonics functions terminology (particularly as applied to stars), the vibratory modes at least for θ and φ are expressed as

$$Y_l^m(\theta/\varphi) = K_0^0 \cdot (\sin \theta \text{ and/or } \cos \theta \text{ terms multiplied by complex exponentials in } \varphi \text{ of the form } e^{\pm im\varphi})$$

where

l (the **spherical-harmonic degree**) is a non-negative integer $l \geq 0$

m (the **azimuthal order**) can take on $2l+1$ integer values in the range $-l$ (ell) to $+l$

K terms are normalization constants.

A third integer parameter n describing the oscillation gives the **radial order**—it determines the harmonic overtone or higher harmonic, where $n = 0$ corresponds to the fundamental. The value of this n parameter is not apparent at the surface and must be inferred indirectly. “The magnitude $|n|$ of the order is then simply the number of radial nodes in the vertical displacement eigenfunction.”²⁷⁰ Thus for the first overtone, $n = 1$ and there is one nodal sphere inside the star.

Physical parameters are given by the real parts of these complex $(a + ib)$ functions.

The radial modes have $l = 0$, $m=0$, while n is variable. The n expresses the higher harmonics (fundamental mode and overtones) and thus the behavior in the radial direction. All radial modes have the form

$$Y_0^0(\theta/\varphi) = K_0^0$$

Stars may exhibit nonradial oscillations, for which $l > 0$ and m can take on values in the range $-l$ (ell) to $+l$. When $l > 0$, the patterns for $m > 0$ represent travelling waves that move across the star parallel to the equator. The **nodal lines** where there is no excursion of radius, define the nodes, as before. They may take the form at the surface of colatitudinal or longitudinal circles. When $m = zero$, the nodal lines are parallel to the equator (have constant θ), the oscillations do not depend on longitude φ , and the pattern is termed **zonal**. When $l = m > 0$, the nodal lines are all great circles passing through the poles (constant φ), the oscillations do not depend on θ , and are referred to as **sectoral**. (*IMA2* p. 503-5 and *here*²⁷¹) For other combinations,

637X/613/1/532/fulltext/ Graph is fig. 1 from the HTML version of this article. (Incidentally, their search for low-dimensional attractors classic for chaotic regimes was not successful.) and

• http://en.wikipedia.org/wiki/Low_dimensional_chaos_in_stellar_pulsations

²⁷⁰ Stellar Oscillations and Helioseismology:

• Christensen-Dalsgaard J and Gough DO, 2001, “On the dipolar f mode of stellar oscillation”, *Monthly Notices of the Royal Astronomical Society*, Volume 326, Issue 3, pages 1115–1121, September 2001. This provides a useful overview of stellar spherical harmonics in general.

• See also <http://arxiv.org/abs/astro-ph/0207403>, “Helioseismology”, *Rev.Mod.Phys.* 74:1073-1129, 2003.

²⁷¹ http://en.wikipedia.org/wiki/Spherical_harmonics

nodal lines divide up the sphere in segments, a pattern referred to as **tesseral** (variously defined, pertaining to tilings with similarly shaped isohedra).

For nonradial pulsations, standing sound waves can propagate horizontally as well as radially.

P-Modes: Sound (acoustic) waves represent perturbations in which pressure P provides the restoring force—these oscillations are called **p-modes**. These modes have both radial and angular **nodes** (angular nodes are defined for θ and/or ϕ). For example, “the p₂ mode with l=4 and m=-3 has two radial nodes between the center and the surface, and its angular pattern has 4 nodal lines, three passing through the poles and one parallel to the equator.” This p₂ mode may be thought of as the nonradial analog to the radial second overtone mode. (*IMA2* p. 505) For nonradial p-modes, most of the motion occurs at and near the surface (*IMA2* fig. 14.17). P-modes tell us mostly about conditions at the surface of a star.

G-Modes: These **Gravity modes** represent gravity waves in which gravity is the restoring force. These are produced by *internal* gravity waves involving sloshing back and forth of stellar gases and related to buoyancy of stellar material. In order to have buoyancy variation, there must be vertical variation, thus $l > 0$. Although peak excursion of nonradial g-modes may be at the surface (*IMA2* fig. 14.18), g-modes exhibit substantial movement in the star’s interior and therefore can provide a view into the very heart of a star. They “are located deep in the solar interior, below the convection zone [of the Sun] (*IMA2* p. 509). (This appears to contradict the figure just cited.)

F-Modes: These modes represent surface gravity waves in which gravity is the restoring force. “The p modes have the higher frequencies which increase as the number of radial nodes increases, whereas the g modes have the lower frequencies, which decrease towards zero as the number of radial nodes increases. In addition, Cowling established numerically for a polytrope of index 3 that for each l there lies between the p and the g modes an *intermediate mode*, with no radial node in the radial displacement eigenfunction. He labeled it the f mode.” Christensen-Dalsgaard tentatively states that one can label modes in a spectrum as follows:

P-Modes have $n > 0$.

F-modes have $n=0$

G modes have $n < 0$ ²⁷²

However, ultimately determining n appears to be more complex than this...

Acoustic Frequency for P-modes: the “angular frequency” of the p-mode is estimated by a computation using the travel time from node to node:

$$S_l = \sqrt{\frac{\gamma P}{\rho}} \frac{\sqrt{l(l+1)}}{r}$$

(I am unclear if this is in radians s⁻¹ or Hz.) The acoustic frequency is large in the deep interior of the star, and decreases with increasing r (*IMA2* p. 506). If there is no rotation, the pulsation period depends only on the number of radial nodes and the value of l. The nominal frequency of the p-wave is an average weighted according to where oscillations are most energetic (the superficial regions). Star rotation introduces a splitting of waves into antegrade and retrograde propagation, and this splitting makes it possible to measure the rotation of the Sun’s interior (*IMA2* p.506).

The Brunt–Väisälä (Buoyancy) frequency for g-modes: This is given as an angular frequency N with a complex formula (*IMA2* p. 508):

$$N = \sqrt{\left(\frac{1}{\gamma P} \frac{dP}{dr} - \frac{1}{\rho} \frac{d\rho}{dr} \right) g} = \sqrt{-Ag}$$

²⁷² Christensen-Dalsgaard J and Gough DO, 2001, op. cit.

The buoyancy frequency is zero at the center of the star (where $g=0$) and at the edges of convection zones (where $A = 0$), and is undefined in a convection zone. The buoyancy frequency of a g-mode is given as the average of N across the star.

Helioseismology and Asteroseismology

This is a rapidly growing highly complex field which I have not explored in any depth and hope to learn more when the opportunity presents.

Adjacent is a computer-generated image showing the pattern of a p-mode solar acoustic oscillation both in the interior and on the surface of the sun. Parameters are $l=20$, $m=16$ and $n=14$. Note that the increase in the speed of sound as waves approach the center of the sun causes a corresponding increase in the acoustic wavelength (apparently referring to the radial or vertical wavelength).²⁷³

Analysis of nonradial pulsations are used in Helioseismology and Asteroseismology. These oscillations were first seen in the Sun in 1962 by Robert Leighton (1919-1997) et al. The solar surface amplitude is usually quite small, with velocities of $< 0.10 \text{ m s}^{-1}$ (measured with Doppler shift of spectral lines) and a luminosity variation $\delta L/L_{\odot} = 10^{-6}$. As many as 10 million modes are simultaneously found in the Sun (by means of Fourier analysis)!

In the Sun, **five-minute oscillations** of periods 3-8 minutes have been found to be p-modes, and are located below the photosphere within the convection zone. The relative power of the p-modes for the Sun peak in the range of frequencies of about 2.5 to 3.5 mHz (milli-Hertz). Many p-modes have been identified.

Differential Rotation has been analyzed with rotational frequency splitting... Below the convection zone, the equatorial and polar rotation rates converge (*IMA2* p. 511). The p-modes have also helped confirm the Sun's surface mass fraction of Helium, $Y = 0.2437$.

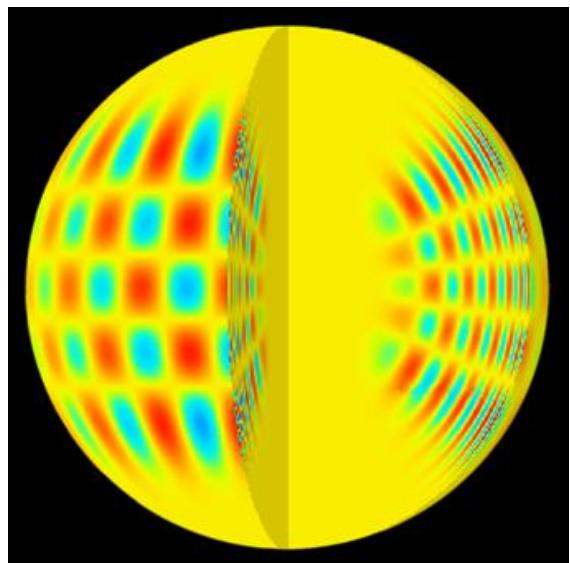
Use of g-mode data for exploring the Sun has so far been unsuccessful (*IMA2* p. 512).

Solar pulsation cannot be explained by the Eddington valve mechanism—although uncertain, it is suspected that solar pulsation draws on the p-modes confined in the convection zone.

The δ Scuti variables²⁷⁴ tend to pulsate in low-overtone radial modes, as well as low order p-modes, over excursions only up to 0.8 mag with a period of minutes of a few to several hours.

SX Phoenicis variable stars exhibit similar mixed mode oscillation with shorter periods of 0.7–1.9 hours.²⁷⁵

The rapidly oscillating Ap stars (**roAp stars**) have strong magnetic fields, pulsation luminosity amplitudes of < 0.016 mag, and apparently exhibit higher order low-degree non-radial p-mode oscillation. Periods of 5.6 to 23.6 minutes are reported. An oblique pulsator model is suggested.²⁷⁶



²⁷³ <http://en.wikipedia.org/wiki/Helioseismology>

²⁷⁴ http://en.wikipedia.org/wiki/Delta_Scuti_variable

²⁷⁵ http://en.wikipedia.org/wiki/SX_Phoenicis_variable

²⁷⁶ http://en.wikipedia.org/wiki/Rapidly_oscillating_Ap_star

The Fate of Massive Stars (Chapter 15)

Post-Main Sequence Massive Stars That Evolve To Supernovae

These are found in the uppermost part of the H-R diagram (on the left except for the RSG), and include **Luminous Blue Variables LBV**, **Wolf-Rayet Stars WR**, **Blue Supergiant stars (BSG, OI or BI**, including Rigel), **Red Supergiant stars (RSG, KI or MI**, including Betelgeuse), and **Of stars** (O supergiants with pronounced emission lines, no example or further details found). Very massive stars are extremely rare—the ratio of $100 M_{\odot}$ to $1 M_{\odot}$ stars is about 1:10⁶, because their lifespans are so short, they are hard to observe, and they just form much less frequently. However, they have a major impact on the ISM by depositing mass and kinetic energy as stellar winds. They can also quench star formation, and their intense UV radiation ionizes gas clouds.

Luminous Blue Variables (LBV)

These are also called **S Doradus variables**, as named after a star in the LMC. They are extremely rare, but well-studied examples include Eta Carinae and P Cygni. Their high luminosity ($> 10^6 L_{\odot}$) and high temperatures (15,000–30,000K) place them in the upper left region of the H-R diagram. Their masses are up to about $125 M_{\odot}$, at or exceeding the Eddington luminosity Limit at which hydrostatic equilibrium is lost due to high radiation pressure, etc.²⁷⁷ This may explain their high mass loss. They may also have atmospheric pulsations instabilities. They may evolve to Wolf-Rayet stars or red supergiants before exploding into supernovae. High rotation velocities of some LBVs may also contribute to the atmospheric instability (IMA2 p. 520).

LBVs are evolved post-main-sequence stars (IMA2 p. 520). They are in an H-R region of instability and may live only short times after turning off the main sequence [MCM: could they never have been on it?]. Modeling has not been perfected, and there are many uncertainties.

Eta Carinae: Eta Carinae (image below²⁷⁸): The massive eruptive variable star system Eta Carinae (η Carinae) at 2300 pc, almost hidden in the center of the Homunculus Nebula, was fitfully variable and then underwent a giant explosion in 1837, increasing from perhaps 2 to 4 mag to -1 mag in 1843 or 1844 (see light curve). This “Great Eruption” period of variable luminosity lasted until 1856. By 1870, it had dropped back to mag 8. It is thought to be a binary,²⁷⁹ consisting of a Luminous Blue Variable (LBV) of about $120 M_{\odot}$ and a much smaller Wolf-Rayet star of about $30 M_{\odot}$. During the Great Eruption period in the 19th century, Eta Carinae probably ejected a mass of $3 M_{\odot}$, visible now as the Homunculus Nebula (occupying much of the image below). The lobes of the “homunculus” are rapidly expanding and contain an elemental signature of the CNO reaction. The LBV has an $T_{\text{eff}} \approx 30,000\text{K}$, so most of its radiation is in the UV. This is scattered and absorbed in the nebula and re-emitted as infrared, giving a deceptively low temperature appearance. “The event that the 19th century astronomers observed was a stellar near-death experience. Scientists call these outbursts **supernova impostor** events, because they appear similar to supernovae but stop just short of destroying their star. The LBV is likely to explode into a full-scale SN in the near future... It is one of the closest stars to Earth that is likely to explode in a supernova in the relatively near future (though in astronomical timescales the “near future” could still be a million years away).²⁸⁰

An interesting partial correlation of eruptive activity and periastron of the binary system is described: “The brief 1838 and 1843 events rose to peak brightness within weeks of periastron passages if the pre-1845 orbital period is ~5% shorter than at present due to the mass loss of the eruption. Each event lasted only

²⁷⁷ http://en.wikipedia.org/wiki/Luminous_blue_variable_star

²⁷⁸ <http://www.spacetelescope.org/images/opo9623a/> 10 June 1996 rotated and cropped by MCM

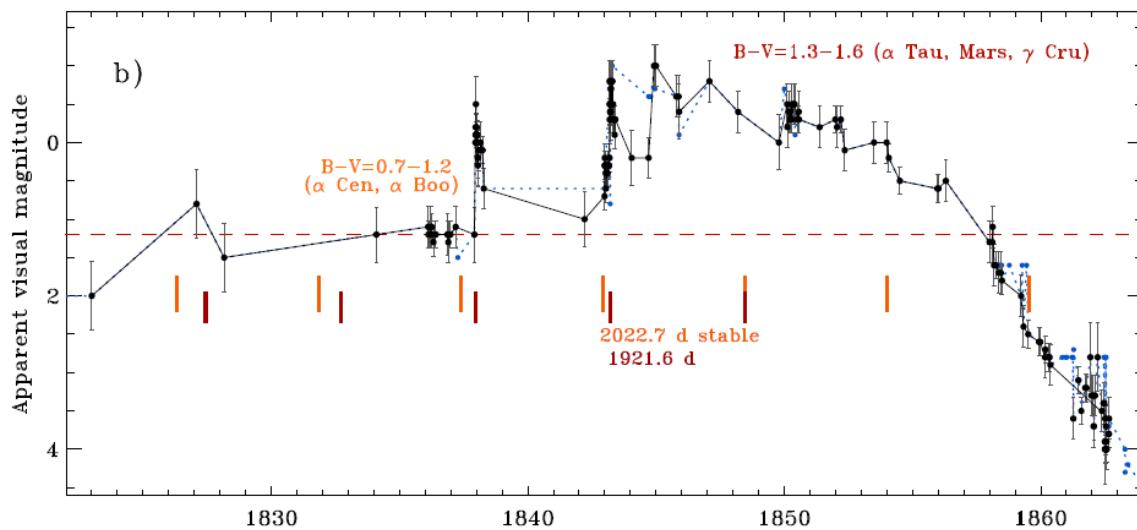
²⁷⁹ http://www.nasa.gov/vision/universe/starsgalaxies/companion_star.html

²⁸⁰ <http://www.spacetelescope.org/images/potw1208a/>

~100 days... The main brightening at the end of 1844 has no conceivable association with periastron, beginning suddenly more than 1.5 yr after periastron.”²⁸¹



Eta Carinae HST composite image, consisting of UV and visible light images from the High Resolution Channel of Hubble’s Advanced Camera for Surveys, 1996.
Distance 7500 ly, ~20 arcseconds across after cropping.²⁸²



This 2010 revision of the historical light curve for Eta Carinae “zooms in on the Great Eruption during 1822–1864. During this time interval, the previous light curve from Frew (2004) is in blue

²⁸¹ Smith N and Frew D, 2010, infra.

²⁸² <http://www.spacetelescope.org/images/potw1208a/>

(points and dotted lines), while the revised light curve... appears as black dots with error bars.

Notes about the apparent color are listed above the light curve.

The orange vertical dashes show predicted times of periastron passage ... whereas
the red hash marks are similar but with a shorter (95%) period before 1848.

The dashed red horizontal line shows the quiescent magnitude of Eta Car as it would appear
with zero bolometric correction.”²⁸³

P Cygni, another LBV, also exhibited periods of marked sudden brightening in 1600 and 1655. Other LBVs include the Pistol Star ($80 - 150 M_{\odot}$), and LBV 1806-20 ($130-200 M_{\odot}$), both of which are highly luminous (see Wikipedia entries for details).

Wolf-Rayet Stars (WR)

These are also evolved massive stars ($>20 M_{\odot}$ initially) with irregular variability, and are closely related to LBVs, though of lower variability. They were discovered in 1867 by French astronomers Charles-Joseph-Etienne Wolf (1827–1918) and Georges Rayet (1839–1906). Around 500 have been cataloged. They have surface temperatures of 25,000K to 100,000K, so are extremely hot. They have strong and very broad emission lines and an unusual spectrum. They lose mass at high rates. Many are rapidly rotating.



Nebula M1-67 surrounding Wolf-Rayet star WR 124 in 2009 (ESA)²⁸⁴,
nebula diameter 1100–1200 arcsec, distance ~5 kpc.

ESA description of this image: “M1-67 is the youngest **wind-nebula** around a Wolf-Rayet star, called WR124, in our Galaxy. These Wolf-Rayet stars start their lives with dozens of times the mass of our Sun, but loose

²⁸³ Smith N, Frew D, “A Revised Historical Light Curve of Eta Carinae and the Timing of Close Periastron Encounters”, 2010, <http://arxiv.org/abs/1010.3719v2>

²⁸⁴ <http://www.eso.org/public/images/wr124/>

most of it through a powerful wind, which is ultimately responsible for the formation of the nebula.... Ten years ago,²⁸⁵ Hubble Space Telescope observations revealed a wealth of small knots and substructures inside the nebula. The same team, led by Cédric Foellmi (ESO), has now used ESO's Very Large Telescope (VLT) to watch how these structures have evolved and what they can teach us about stellar winds, their chemistry, and how they mix with the surrounding interstellar medium, before the star will eventually blow everything away in a fiery supernova explosion... The image is based on FORS1 data obtained by the Paranal Science team with the VLT through 2 wide (B and V) and 3 narrow-band filters.²⁸⁶

There are 3 categories of WR—**WN**, **WC**, and **WO**—based on content of N, C, O as well as He and H and highly ionized species. Numerical subdivisions (such as WC9) are used to express degrees of ionization (*IMA2* p. 522). The predominance of N in WN and the trends in relative abundance from WN to WO WRs correlate with mass loss, starting with the H envelope, and with convection from the CNO processed core, exposing the elements nucleosynthesized in the core. “If the star survives long enough, mass loss will eventually strip away all but the oxygen component of the triple-alpha ash [i.e., ash of C and O].” (*IMA2* p. 522)

Evolutionary Scheme for Massive Stars

In one version of the scheme first advanced by Peter Conti in 1976, massive stars end their lives as supernovae—the sequences based on stellar mass are as follows:²⁸⁷

$M > 85 M_{\odot}$:	$O \rightarrow LBV \rightarrow WN \rightarrow WC \rightarrow SN$
$40 M_{\odot} < M < 85 M_{\odot}$:	$O \rightarrow WN \rightarrow WC \rightarrow SN$
$25 M_{\odot} < M < 40 M_{\odot}$:	$O \rightarrow RSG \rightarrow WN \rightarrow WC \rightarrow SN$
$20 M_{\odot} < M < 25 M_{\odot}$:	$O \rightarrow RSG \rightarrow WN \rightarrow SN$
$10 M_{\odot} < M < 20 M_{\odot}$:	$O \rightarrow RSG \rightarrow BSG \rightarrow SN$

IMA2 (p. 522) lists this sequence slightly differently, as follows:

$M > 85 M_{\odot}$:	$O \rightarrow Of \rightarrow LBV \rightarrow WN \rightarrow WC \rightarrow SN$
$40 M_{\odot} < M < 85 M_{\odot}$:	$O \rightarrow Of \rightarrow WN \rightarrow WC \rightarrow SN$
$25 M_{\odot} < M < 40 M_{\odot}$:	$O \rightarrow RSG \rightarrow WN \rightarrow WC \rightarrow SN$
$20 M_{\odot} < M < 25 M_{\odot}$:	$O \rightarrow RSG \rightarrow WN \rightarrow SN$
$10 M_{\odot} < M < 20 M_{\odot}$:	$O \rightarrow RSG \rightarrow BSG \rightarrow SN$

Note that SN do not occur at progenitor masses $M_{\odot} < 8$ (range about $6 M_{\odot}$ to $9.5 M_{\odot}$ for this lower limit for SNe, upper limit for WD).²⁸⁸ These masses are probably ZAMS masses, but I am not certain.

Woosley et al 2002 state, “Unlike lower-mass stars ($M \lesssim 8 M_{\odot}$), however, no point is ever reached at which a massive star can be fully supported by electron degeneracy. Instead, the center evolves to ever higher temperatures, fusing ever heavier elements until a core of iron is produced. The collapse of this iron core to a neutron star releases an enormous amount of energy, a tiny fraction of which is sufficient to explode the star as a supernova... ”²⁸⁹

²⁸⁵ <http://hubblesite.org/newscenter/archive/releases/1998/38/image/a/> (also shown at *IMA2* p. 521)

²⁸⁶ <http://www.eso.org/public/images/wr124/>

²⁸⁷ Massey P, “Massive Stars in the Local Group: Implications for Stellar Evolution and Star Formation”, *Astronomy and Astrophysics*, Vol. 41: 15–56, 2003

²⁸⁸ Williams KA, et al, “Probing The Lower Mass Limit for Supernova Progenitors and the High-Mass End of the Initial-Final Mass Relation from White Dwarfs in the Open Cluster M35 (NGC 2168)”, *The Astrophysical Journal*, 693:355–369, 2009 March 1, doi:10.1088/0004-637X/693/1/355.

²⁸⁹ Woosley SE, et al, “The evolution and explosion of massive stars”, *Rev. Mod. Phys.* 74, 1015–1071 (2002)

The evolutionary tracks for massive stars ending in SNe are shown as follows (Meynet and Maeder 2003). These tracks begin on the left, near the mass label, probably at the ZAMS. The tracks for rotating $25 M_{\odot}$ and $40 M_{\odot}$ return on themselves in a final blueward leg of a loop. (The track for $20 M_{\odot}$ also has a blueward loop, not shown here, *IMA2* p. 537.) Mass loss is significant and is included in these models. Rotation is also important because it drives internal mixing and enhancing mass loss (and the tracks show a corresponding greater luminosity for rotation). Note that the most massive stars never become red supergiants. (*IMA2* p. 523 and *here*²⁹⁰):

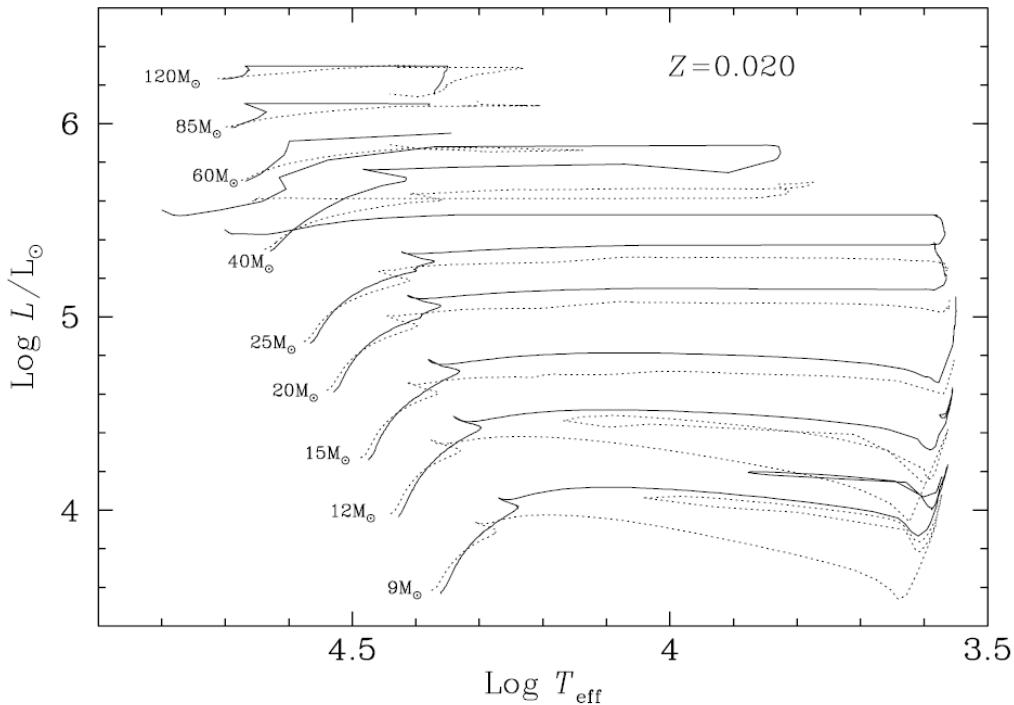


Fig. 5. Evolutionary tracks for non-rotating (dotted lines) and rotating (continuous lines) models for a metallicity $Z = 0.020$. The rotating models have an initial velocity v_{ini} of 300 km s^{-1} , which corresponds to an average velocity during the MS phase of about 180 to 240 km s^{-1}

The Humphreys-Davidson Limit

Roberta Humphreys and Kris Davidson were the first to observe that above a certain mass of about $40 M_{\odot}$, an upper luminosity cutoff occurs, and by implication an upper limit to the masses of stars that could evolve to become cooler red supergiants. “It became apparent that the upper luminosity limit varies non-linearly with temperature for hot stars. For these early type [massive] stars the maximum luminosity decreases with decreasing effective temperature, this downward trend ceasing at around $T_{\text{eff}} = 10^4 \text{ K}$, after which the maximum luminosity remains more or less constant. Humphreys and Davidson (1979) noted this boundary and suggested that an instability leading to rapid and unsteady mass loss set in there. The observed boundary could, in principle, be explained by steady mass loss, but only if the loss were much higher than that which is actually observed for objects near the boundary (typically $10^{-5} M_{\odot}$ per year), so it is the ‘unsteady’ mass loss in sporadic eruptions which is considered to be of foremost importance.”²⁹¹

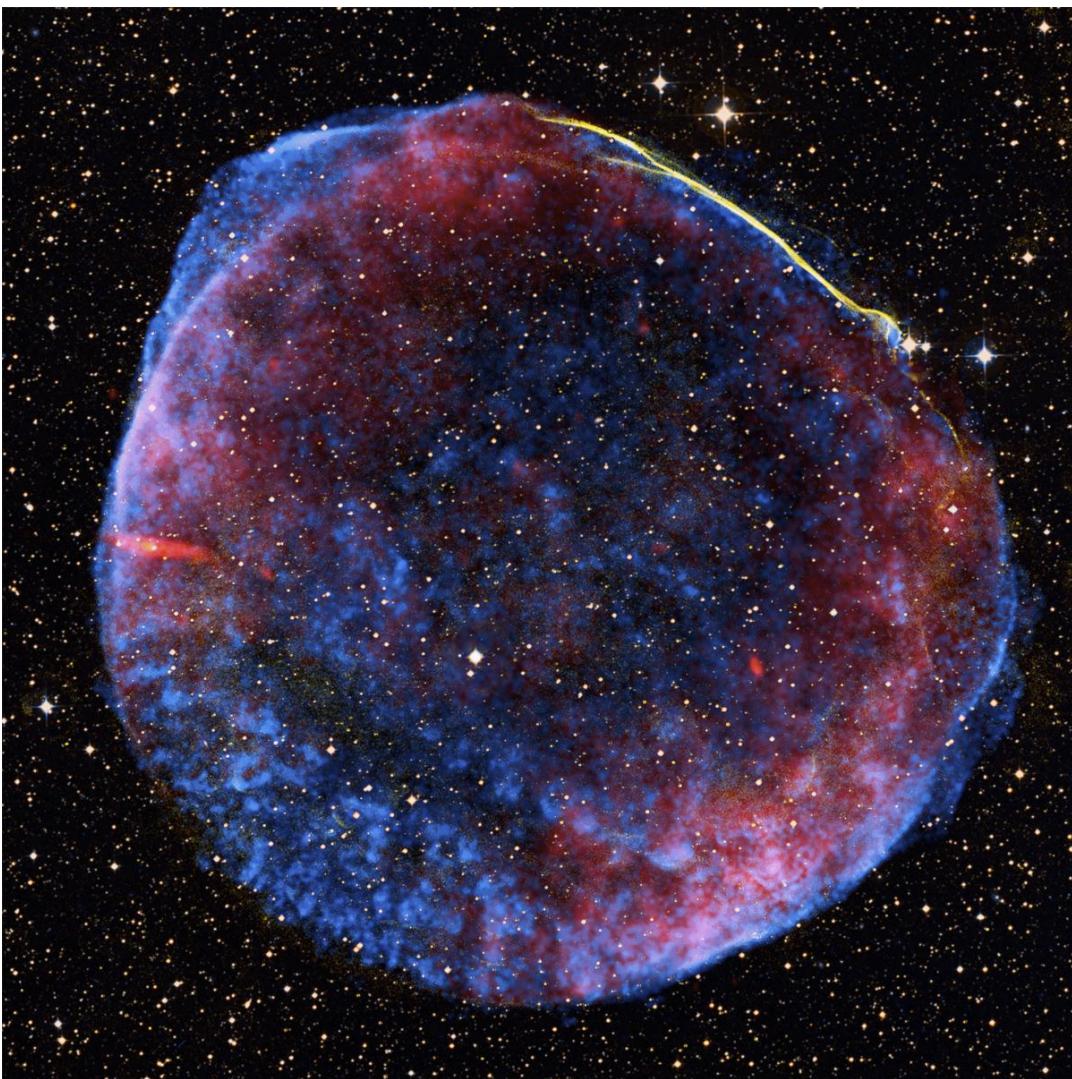
²⁹⁰ Meynet G and Maeder A, “Stellar evolution with rotation - X. Wolf-Rayet star populations at solar metallicity” *Astronomy and Astrophysics*, Vol. 404, No. 3 (June IV 2003) p. 975, including graph shown
DOI: 10.1051/0004-6361:20030512

²⁹¹ <http://www.peripatus.gen.nz/astronomy/HerRusDia.html>

Supernovae (Supernovas)

Important Historical Supernovae

SN 1006: The brightest SN in history was the **SN of 1006**, which reached an apparent visual magnitude of about $V = m_V = -9$, and left the following supernova remnant (SNR) (see image below). It was witnessed by naked eye observers in China, Egypt, Iraq, Japan, Switzerland, and possibly North America.²⁹² The Chandra XRO page states, “This is a composite image of the SN 1006 supernova remnant, which is located about 7000 light years from Earth. Shown here are X-ray data from NASA’s Chandra X-ray Observatory (blue), optical data from the University of Michigan’s 0.9 meter Curtis Schmidt telescope at the NSF’s Cerro Tololo Inter-American Observatory (CTIO; yellow) and the Digitized Sky Survey (orange and light blue), plus radio data from the NRAO’s Very Large Array and Green Bank Telescope (VLA/GBT; red)... This combined study of the Chandra, CTIO and VLA/GBT observations shows new evidence for the acceleration of charged particles to high energies in supernova shockwaves. An accompanying Hubble Space Telescope image of SN 1006²⁹³ shows a close-up of the region on the upper right of the supernova remnant. The twisting ribbon of light seen by Hubble reveals where the expanding blast wave is sweeping into very tenuous surrounding gas.”



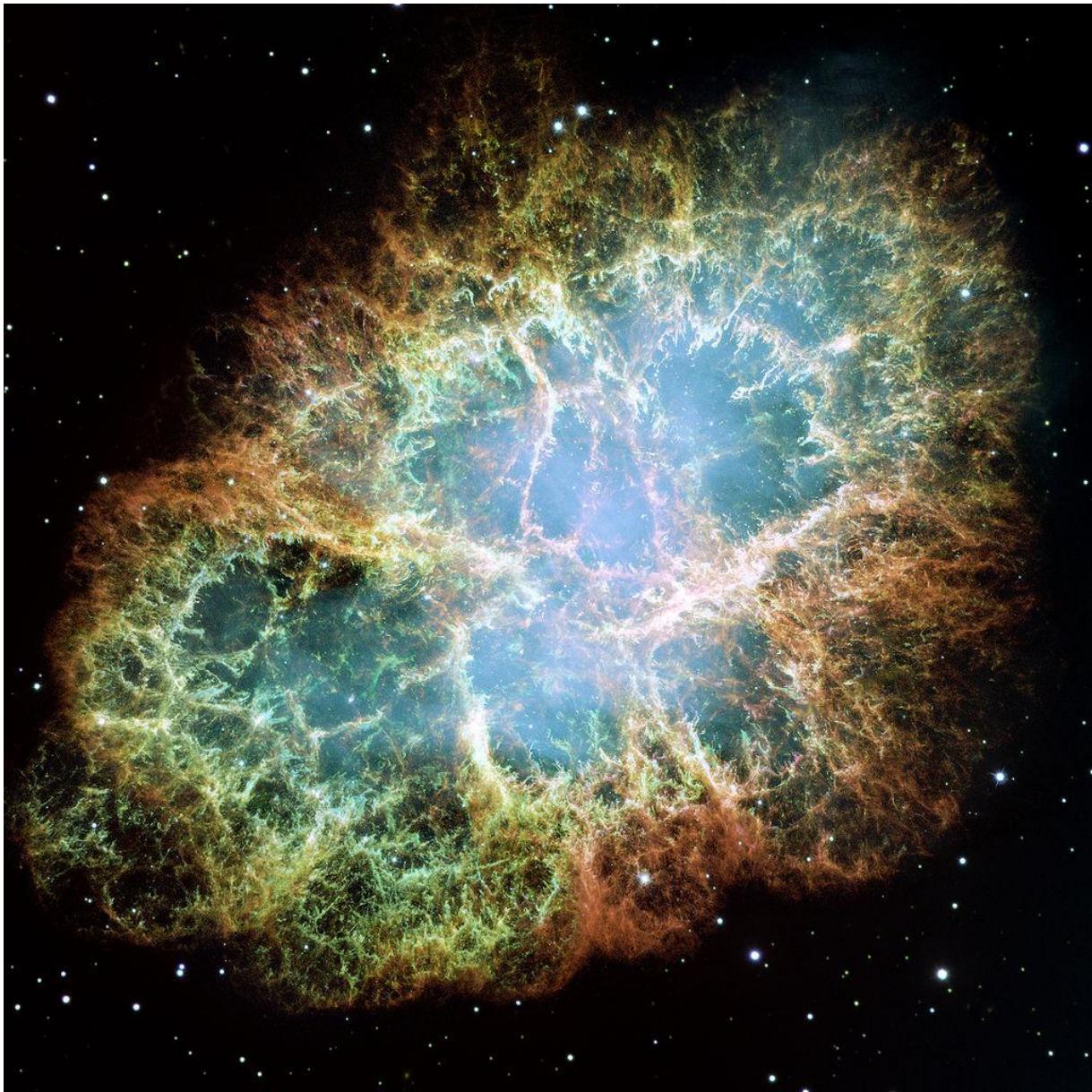
SN 1006 remnant at distance 7000 ly = ~2.2 kpc, 55 arcmin diameter ~20 pc.
(combined from X-ray, visible, and radio, 2008)²⁹⁴

²⁹² http://en.wikipedia.org/wiki/SN_1006

²⁹³ http://hubblesite.org/hubble_discoveries/science_year_in_review/pdf/2008/supernova_remnant_sn_1006.pdf

²⁹⁴ <http://chandra.harvard.edu/photo/2008/sn1006c/> released 2008, images detected 2003

SN 1054: The SN of 1054 (a Type II) gave rise to a well-known SNR, the **Crab Nebula (M1)**.²⁹⁵ HST states, “The newly composed image [released in 2005] was assembled from 24 individual Wide Field and Planetary Camera 2 exposures taken in October 1999, January 2000, and December 2000. The colors in the image indicate the different elements that were expelled during the explosion. Blue in the filaments in the outer part of the nebula represents neutral oxygen, green is singly-ionized sulfur, and red indicates doubly-ionized oxygen.” Pulsar PSR B0531+21 is the central star in the Crab Nebula.



Crab Nebula, at distance $6500 \text{ ly} = 2 \text{ kpc}$, width $6 \text{ arcmin} = 12 \text{ ly} = 3.7 \text{ pc}$
(HST, composite image released 2005)²⁹⁶

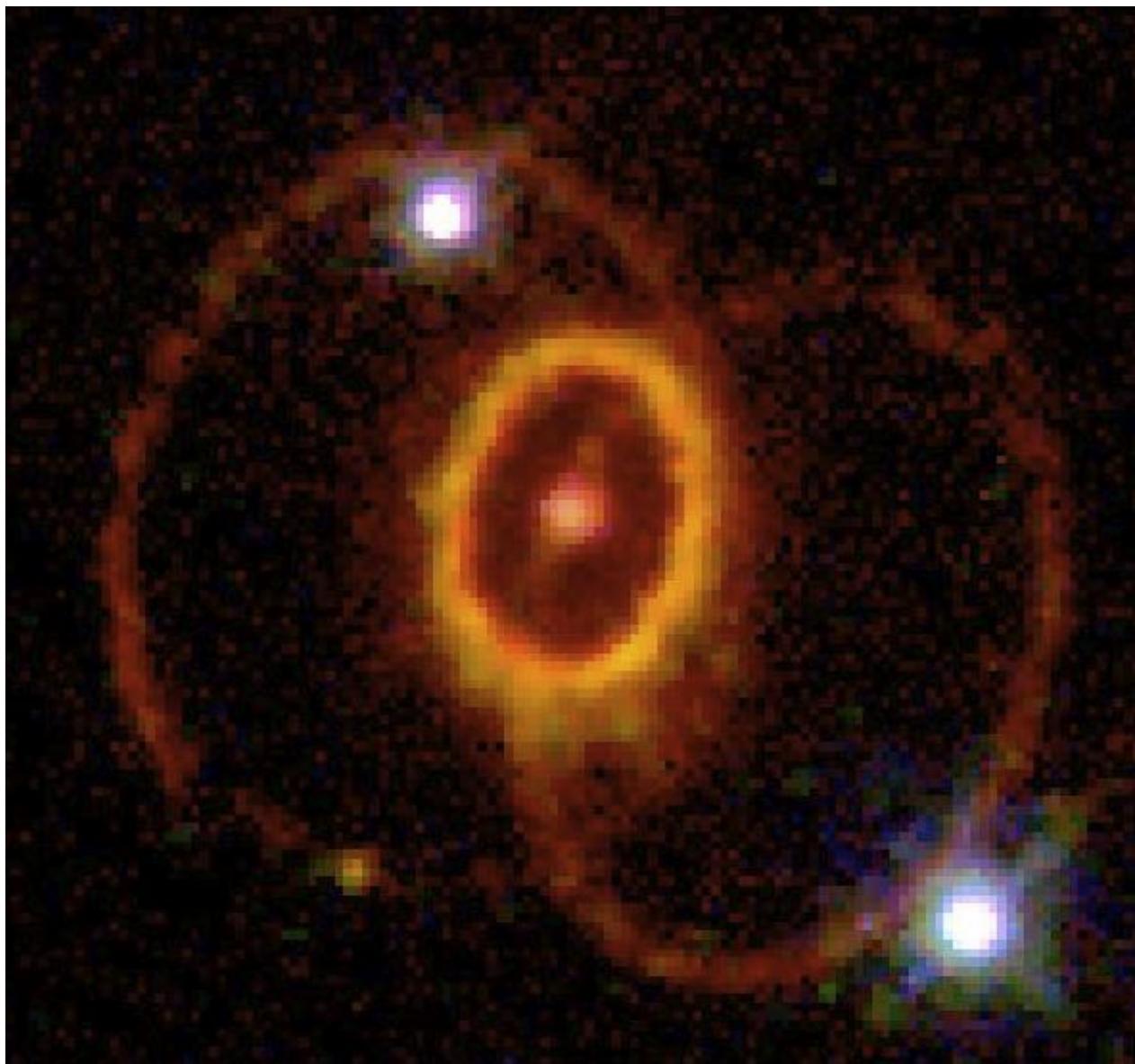
SN 1572, Tycho's SN: Tycho Brahe witnessed naked-eye SN in 1572—probably a SN Type Ia. It arose in the MW.

SN 1604, Kepler's SN: Tycho's pupil Johannes Kepler also saw a naked-eye SN, namely SN 1604 in 1604—probably also a SN Type Ia and arising in the MW. It was, as of May 2012, the last SN observed that arose in the MW and the last naked-eye SN until SN 1987A.

²⁹⁵ <http://hubblesite.org/newscenter/archive/releases/2005/37/>

²⁹⁶ <http://hubblesite.org/newscenter/archive/releases/2005/37/>

SN 1987A: The SN 1987A, a Type II-P, is located at 51.5 kpc in the nearby LMC (50 kpc), thus not in the MW, and arose from a blue supergiant star (BSG). It is the first naked-eye SN since Kepler's SN 1604. A 1995 HST image of the region of the SN (detected in 1994), follows:



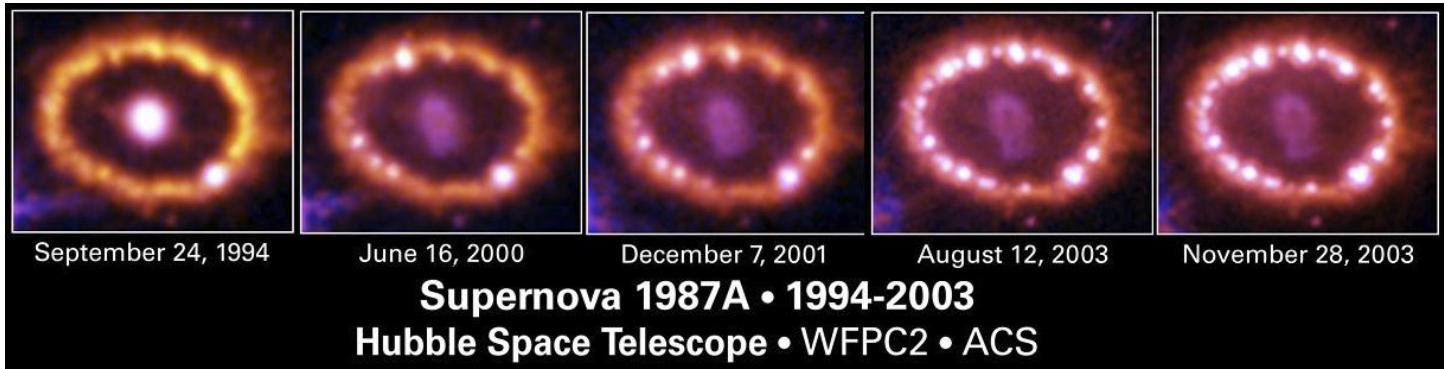
Rings associated with SN 1987A (image detected 1994, released 1995)²⁹⁷
Distance 51.4 kpc

This Type II-P SN is providing a trove of very detailed information about the evolution of SNRs. The remnants of this SN are unusual and merit greater study than I have so far given them.

The substance of the bright **inner circumstellar ring**, which lies in the equatorial plane [of the binary system] that contains the center of the SN (*IMA2* p. 538), is believed to have been ejected by the progenitor BSG star (Sanduleak -69° 202) approximately 20,000 years ago. This equatorial plane is inclined 43° with respect to the line of sight.²⁹⁸ This inner ring began exhibiting variable clumps of brightening in 1996 (*IMA2* p. 539, see image below), as the faster expanding SN 1987A shock front overtook and collided with the older ring. The visible light glow from this ring arises from O III emissions (*IMA2* p. 538).

²⁹⁷ <http://hubblesite.org/newscenter/archive/releases/1995/49/image/a/> image detected in 1994, released 1995.

²⁹⁸ Larsson J, et al, "X-ray illumination of the ejecta of supernova 1987A", *Nature* 474, 484–486, 23 June 2011, doi:10.1038/nature10090.



Progressive development of knots or clumps of brightening in the old inner ring surrounding SN 1987A, images excerpted by MCM from HST images.²⁹⁹ Ring is ~1 ly across.
Constant star ~4 o'clock is unrelated.

The two **larger outer rings** lie in different planes, each parallel to the equatorial plane but in front of and behind the equatorial plane, resp., and appear offset from the center of the SN site. *IMA2 p. 538* offers an interesting hypothesis that these may be “painted” on formerly ejected hourglass-shaped mass by narrowly collimated jets of radiation emanating from the wobbling companion star, but this is not universally accepted. Alternatively, “A possible explanation for the formation of the [outer] rings is that they were formed by interaction of the [stellar] wind from the progenitor star [a BSG] with gas that was released from the star at earlier stages.”³⁰⁰ “Many authors showed that the rings of SNR1987A coincide with the rim of a structure having the shape of an hourglass. The hourglass may be a Strömgren shell distorted by the inhomogeneity of the gas or of the radiation of the star.”³⁰¹ *Here*³⁰² is an animated visualization of the distribution of the 3 rings and the associated partly-hourglass-like surface in question. Obviously, this topic will undergo continued development.

The outer rings apparently do not consist of matter ejected in 1987, but rather may have been ejected ~20,000 years ago. Therefore, they may not be true SNRs. Further comments about this SN will be made below.

This supernova was **subluminous** compared to other Type II SNe. Its peak brightness was $M_{bol} = -15.5$ (*IMA2 p. 536*), whereas typical Type IIs attain $M_{bol} = -18$. Analysis shows that this is because its progenitor (with mass $\sim 20 M_{\odot}$) had evolved from a RSG to a BSG along a blueward loop, attaining greater density and a deeper potential well than the usual RSG.

This SN allowed detection of fully resolved **gamma rays** from the decay of $^{56}_{28}\text{Co}$, namely the emissions at 1238 keV³⁰³ and 847 keV, as well as 2599 keV.³⁰⁴

Neutrinos from this SN were detected at several sites, arriving 3 hours before the photon burst arrived. The plausible neutrinos detected numbered only 24. According to models, the neutrinos originated and escaped before the shock wave reached the surface of the star to trigger the release of photons. The neutrino speed was deduced to be very close to the speed of light, $(1 - 10^{-8})c$. (*IMA2 p. 540*).

The expected compact stellar remnant at the core of the SN has never been detected, nor has a companion star surviving after the SN.

Light echoes were observed in 1988 and later for this and other SNe, some enhanced by difference imaging.³⁰⁵

²⁹⁹ <http://hubblesite.org/newscenter/archive/releases/2004/09/image/b/>

³⁰⁰ Tziamtzis A, et al, “The outer rings of SN1987A”, <http://arxiv.org/abs/1008.3387v3>, 2010, to be published in *Astronomy & Astrophysics*.

³⁰¹ http://en.wikipedia.org/wiki/Supernova_1987a

³⁰² <http://i.ytimg.com/vi/wez3ElEtPT4/default.jpg>, click on “Supernova SN1987a rendered with glnemo2”; see also Morris T, Podsiadlowski Ph, “The triple-ring nebula around SN 1987A: fingerprint of a binary merger.” 2007, *Science*, 315, 1103, including “Supporting Online Material” DOI: 10.1126/science.1136351, which has 2 additional movies.

³⁰³ Mahoney WA et al, “Gamma-ray observations of Co-56 in SN 1987A”, *Astrophysical Journal, Part 2 - Letters*, vol. 334, Nov. 15, 1988, p. L81-L85. Detected 1238 keV line but not 847 keV line.

³⁰⁴ Tueller J, et al, “Observations of gamma-ray line profiles from SN 1987A”, *Astrophysical Journal, Part 2 - Letters*, vol. 351, March 10, 1990, p. L41-L44. Included lines at 1238 and 2599 keV.

The Classification of Supernovae

Supernovae are rare, occurring about once in 100 years in a galaxy. Based on spectra and light curves, they are classified as follows:

Type I: lacks H lines in the spectrum (star was stripped of its H envelope)

Type Ia: have strong Si II line at 615 nm near peak light. Found in all types of galaxies.

Type Ib: no strong Si II line at 615 nm, strong He I lines at 586.6 nm, etc. Only in spirals.
May be from WN WR (Wolf-Rayet) stars.

Type Ic: no strong Si II line at 615 nm, weak or no He lines. Only in spirals. May be from WC WR stars.

Type II: have strong H lines, such as H_α and H_β , more common than Ib or Ic, usually red supergiant

Type II-P: Reaches a (sloping) plateau during 30–80 d after peak.

Type II-L: Linear decrease in light curve, no plateau.

Type II-N: have some narrow lines (this type is not mentioned in *IMA2*, see *here*).³⁰⁶

Type I: Type Ia's release about $1-2 \times 10^{44}$ J and have absolute magnitudes $\langle M_B \rangle \cong \langle M_V \rangle \cong -19.3 \pm 0.3$. Types Ib and Ic average 1.5 to 2 mag less bright compared to Type Ia. Type Ib's release 0.1 to about 1×10^{44} J and have absolute magnitude from -15 to -17 . Type Ic's release 0.1 to about 5×10^{44} J and have absolute magnitude from -16 to -22 , thus often brighter than Ia's or Ib's. Types Ib and Ic are found in spirals near active star formation, suggesting they arise in massive short-lived stars. Type I's decline at ~ 0.065 mag d $^{-1}$ at 20 d, the rate of decline slowing at 50 days.³⁰⁷

Type II: Type II SNe have a rapid rise in luminosity. They may exhibit P Cygni profile (previously discussed). They release 0.1 to 5×10^{44} J (*IMA2* p. 528 says typically 10^{46} J), thus often more than Type I's. The energy released is 1% in the form of particle KE, 0.01% as photons, and about 99% as neutrino energy. The energy released is not due to fusion reactions, but to release of gravitational potential energy. Type II-P's release 0.1 to 1×10^{44} J and have absolute magnitude -14 to -16 . Type II-L's release 1×10^{44} J and have absolute magnitude -17 . Type II-N's release 1 to 5×10^{44} J and have absolute magnitude as high as -22 . (These energy release estimates are lower than in *IMA2*).³⁰⁸

Pair instability SNe: This is proposed mechanism in *luminous supernovas* (those with absolute magnitudes > -21), arising in very large stars with very large cores, for which the type for many is said to be Ic, and having absolute magnitudes of up to greater than -22 (e.g., PTF09cnd): "...Extremely massive stars with $M_{\text{initial}} \geq 140 M_\odot$ (if such exist) develop oxygen cores with masses, M_{core} , that exceed $50 M_\odot$, where high temperatures are reached at relatively low densities. Conversion of energetic, pressure-supporting photons into **electron-positron pairs** occurs before oxygen ignition and leads to a violent contraction [due to sudden drop in radiation pressure] which triggers a nuclear explosion that unbinds the star in a pair-instability supernova. Transitional objects with $100 M_\odot < M_{\text{initial}} < 140 M_\odot$ may end up as iron-core-collapse supernovae following violent mass ejections, perhaps as a result of brief episodes of pair instability, and may already have been identified. Here we report observations of supernova **SN2007bi**, a luminous, slowly evolving object located within a dwarf galaxy. We estimate the exploding core mass to be $M_{\text{core}} \approx 100 M_\odot$, in which case theory unambiguously predicts a pair-instability supernova. We show that $> 3 M_\odot$ of radioactive ^{56}Ni was synthesized during the explosion and that our observations are well fitted by models of pair-instability supernovae..."³⁰⁹

³⁰⁵ Light Echoes from Supernova 1987A:

- <http://www.eso.org/public/news/eso8802/>
- <http://apod.nasa.gov/apod/ap060125.html>
- <http://universe.gsfc.nasa.gov/conferences/supernova1987a/Crotts.pdf>

³⁰⁶ <http://en.wikipedia.org/wiki/Supernova>

³⁰⁷ *ibid.*

³⁰⁸ *ibid.*

³⁰⁹ Pair-Instability SNe:

- Gal-Yam A., et al, "Supernova 2007bi as a pair-instability explosion", *Nature* Vol 462, 3 December 2009, doi:10.1038/nature08579, <http://www.nature.com/nature/journal/v462/n7273/abs/nature08579.html>
- Quimby RN, et al, "Hydrogen-poor superluminous stellar explosions", *Nature* 23 June 2011 Vol 474 , p. 487, doi:10.1038/nature10095
- Gai-Yam A, "Super Supernovae", *Scientific American* 306, 44 - 49 (2012)
doi:10.1038/scientificamerican0612-44

Core Collapse SNe

These occur in Types Ib, Ic, and II SNe. (Type Ia's are fundamentally different, and discussed in Chapter 18.) The mechanism of core-collapse SN are similar in these 3 types. In stars $> \sim 8 M_{\odot}$, the final stage of stellar evolution is not a planetary nebula but a SN.

The details of the catastrophic collapse of the iron core leading to a SN (see *IMA2* p. 530 to 534) are complex and not fully understood. Computer simulation is difficult and strongly sensitive to initial conditions

Briefly, the He-burning core builds up C and O, eventually igniting C burning. An onion-like structure of concentric shells develops in the deep interior or core—for each element involved, an non-burning layer enriched in an element overlies a burning enriched layer of the same element. From the core outward, the layer pairs are enriched in Fe (the Fe core), Si, O, C, He, and, in the envelope, H/He. Elements heavier than ^{56}Fe are created **endothermically** and cannot contribute to the star's luminosity. Each of the elements named, other than Fe, i.e., from He to Si, undergoes fusion and “burns” in a shell of thermonuclear reactions (“oxygen burning”, “carbon burning”, etc.) Progressing inward, each of the successive burning layers, generates less energy than the previous burning layer outside it, and is of shorter duration than the previous. Core H burns 10^7 y, core He burns 10^6 y, core C burns 300 y, core O burns 200 d, and core Si burning lasts only 2 days. The high core temperature causes energetic photons to begin to destroy heavy nuclei (by **photodisintegration**), leading to conversion of some of the Fe back to He and H ions. This is highly endothermic, and removes thermal energy from the gas, reducing its pressure, and this pressure is needed to support the core against collapse. The core has a mass of $1.3 M_{\odot}$ to $2.5 M_{\odot}$, depending on the star's starting (ZAMS) mass.. At this point, the T is about 8×10^9 K. The free electrons that contribute to needed degeneracy pressure are captured by the protons and heavier nuclei (by *electron capture*). The electron capture reaction is a beta decay, and releases neutrons and much energy in the form of neutrinos that escape. The photodisintegration and electron captures have the combined effect of greatly and suddenly **reducing the electron degeneracy pressure**. The inner **catastrophic iron core suddenly collapses** at speeds of up to 70,000 km s⁻¹ (0.2c), and the core is strongly compressed (details get a little confusing). The outer layers are no longer supported by the core. The inner core suddenly stiffens up due to Pauli exclusion principle applying to compression of neutrons (which are also Fermions obeying Fermi-Dirac statistics), and this causes a bounce or rebound, sending pressure and shock waves outward toward the infalling outer core. Even neutrinos have difficulty escaping in the high density that results, causing heating in this layer. Eventually, if the star mass and other parameters are right, the shock wave reaches the surface and blows off the envelope. (Shock waves that are dissipated before they reach the surface abort the formation of a SN.) The envelope is driven outward by the shock. The KE of the ejected material is about 10^{44} J, the photons carry away 10^{42} J (causing a peak luminosity of $10^9 L_{\odot}$), and the release of neutrinos carries off about 10^{46} W (!)

In summary, the key events are the catastrophic collapse of the iron core, the generation of a shock wave, and the ejection of the star's envelope.

The differences among Types Ib, Ic, and II arise from differences in envelope makeup and the amount of radioactive material synthesized. They may also relate to which type of WR star precedes them.

Stellar Remnants

After a core-collapse SN, the inner core may become a **neutron star** (for stars with $M_{\text{ZAMS}} < 25 M_{\odot}$) or a **black hole** (for larger stars). (*IMA2* p. 534) However, the star may simply be blown apart without leaving a stellar remnant. Type Ia, though leaving a supernova remnant, do not leave a neutron star stellar remnant,³¹⁰ whereas core-collapse SN do—but I am unsure whether there are any Type Ia exceptions to this.)

Ionization and Recombination

The light curves exhibited after the peak of a SN explosion reflect the **recombination** of ions and electrons that were split apart (ionized) by the shock. The increased ionization causes greater opacity of the outer envelope, preventing initial escape of photons. The gradual recombination of ions and electrons is

³¹⁰ http://chandra.harvard.edu/xray_sources/supernovas.html

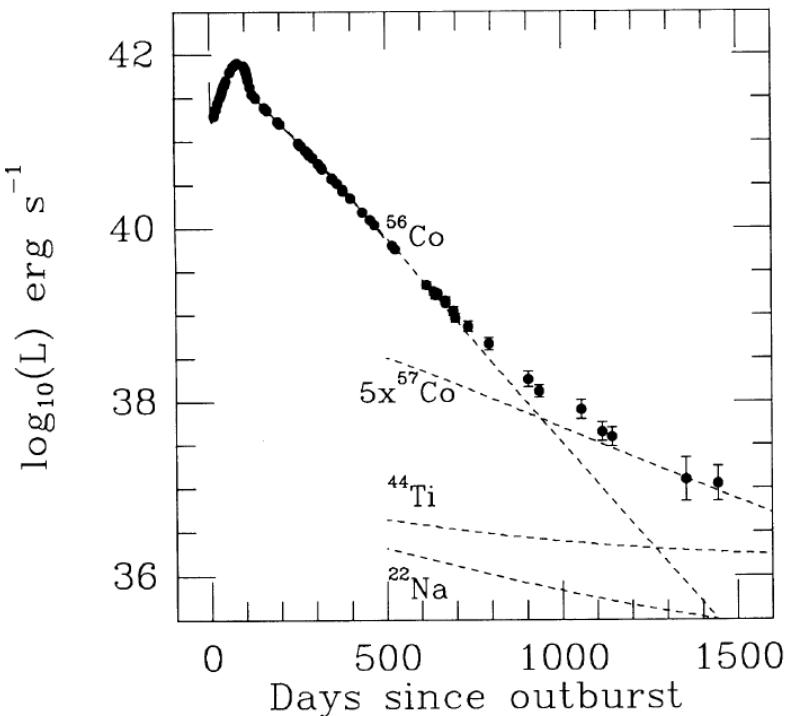
accompanied by increasing transparency and gradual emission of energy from this recombination at an effective temperature of about 5000K. (IMA2 p. 534)

Generation of Radioisotopes

In addition to energy release from recombination, the presence of recently synthesized short-lived **radioisotopes**, synthesized during the passage of the shock front through the star, also contributes to the light curve. The plateau seen with a Type IIP SN is in part attributable to energy deposited in the ejecta from radioactive decay of ^{56}Ni , which has a half-life $t_{1/2}$ of only 6.1 days. (IMA2 p. 534)

This graph³¹¹ depicts the light curve observed for SN 1987A extending to day 1444. The time to maximum light was slower than usual for Type IIs, taking 80 days to attain a somewhat weak peak of $M_{\text{bol}} = -15.5$ (IMA2 p. 536). The solid upper line from day 1 to 500 represents bolometric flux based on the U to M bandpass integration. The dotted lines represent estimates of energy deposition from the four

radionuclides shown according to the expected model proportions (except that ^{57}Co is increased to 5 times that expected based on the solar ratio of Fe-57/Fe-56, in order to better match the observations). Note that the article quotes decay times in terms of e-folding decay times, not half-lives. After the rapid initial falloff just past the peak (which is contributed to by rapid decay of ^{56}Ni ($t_{1/2} = 6.1$ days)), the rate of decline slows (reflecting relative depletion of the ^{56}Ni). The curve now follows the decay curve of ^{56}Co ($t_{1/2} = 77.2$ days). Eventually, at about 1000 days, this isotope also becomes depleted, and a further slowing of the decay occurs, now probably following the radioisotope decay curve for ^{57}Co ($t_{1/2} = 272$ days). See the article for estimates of the amounts of these radionuclides generated. Characteristic X-ray and gamma ray emissions from the decay chain of ^{57}Co were measured for the first time after SN 1987A, namely lines at 847 keV and 1238 keV (cited earlier).



Supernova Remnants (SNRs)

These are common, and include the Crab Nebula from SN 1054, the SN 1006 SNR, the SN 1572 SNR, and the SN 1604 SNR. The ejected material is travelling at up to 0.1c (30,000 km s⁻¹), and a strong shock wave is formed defining the leading edge of the SNR as it excites and ionized the ISM. Some of the radiation may be in the form of synchrotron radiation indicating relativistic electrons spiraling around magnetic field lines, fed by ongoing energy emission from a pulsar stellar remnant. The term SNR applies to ejected material, not to a stellar SN remnant such as a NS.

Chemical Abundances and SN Nucleosynthesis

A chart of solar system number abundances of the elements has been previously provided. The textbook (IMA2 p. 541) shows a similarly plotted graph for the relative abundance of the Sun's photosphere. H and He are much more abundant there, and are found in a 10:1 ratio (by numbers). Li, Be, and B are underabundant and can be easily destroyed in collisions with protons. Destruction of Be requires a higher temperature than Li. The Sun's convection zone transports surface Li, Be, and B into the interior. The Sun has an unexpectedly low Li, a finding that is inconsistent with current solar models. The elements C, O, Ne,

³¹¹ Suntzeff NB, et al, "The energy sources powering the late-time bolometric evolution of SN 1987A", *Astrophysical Journal, Part 2 - Letters*, vol. 384, Jan. 10, 1992, p. L33-L36,

<http://dx.doi.org/10.1086/186256>

Mg, and Si are relatively more abundant, due to their nuclei being multiples of alpha-particles (modified from *IMA2* p. 542 by excluding N). Core-collapse SN generate much O (whereas Type Ia SN generate much Fe).

S-Process and R-Process Nucleosynthesis

Elements with higher Z are harder to penetrate by protons due to a high Coulomb potential barrier, but neutrons suffer no such repulsive barrier. Once a neutron combined with a nucleus, raising the atomic mass number A to A + 1, the new nucleus may undergo beta decay. This process emits an electron, a neutrino, and a gamma ray, leading to an increase of atomic number from Z to Z+1.

When beta decays are relatively rapid relative to N captures and the neutron captures are rate limiting, the process is called an **S-process (slow-neutron-capture process)**, and these tend to be present in normal stellar nucleosynthesis and to yield stable nuclei. The S-process produces approximately half of the isotopes of the elements heavier than iron (the other half from R-Processes).

When beta decays are relatively slow and rate limiting compared to more rapid neutron captures (such as immediately after a core-collapse SN, when there is an extremely high neutron flux), the process is called an **R-process (rapid-neutron-capture process)**. These occur commonly in SN reactions, and result in relatively N-rich nuclei.

The two types of processes therefore play important roles in chemical evolution. They account for the abundances of elements with A > 60.

Gamma-Ray Bursts GRBs and Associated X-ray Emissions

GRBs were first noted with the Vela satellites launched in the 1960s to monitor the compliance of Soviet and other nuclear nations with the 1963 Partial Test Ban Treaty through detection of gamma rays. In 1967, the Vela 4 and Vela 3 satellites first detected a flash of gamma radiation unlike a nuclear weapons signature... As additional Vela satellites were launched with better instruments, the Los Alamos team continued to detect inexplicable gamma-ray bursts. By analyzing the different arrival times of the bursts as detected by different satellites, the team was able to rule out a terrestrial or solar origin. The discovery was declassified and published in 1973.³¹² GRBs are not the same as Soft Gamma Repeaters SGRs (described later).

Currently active NASA-related satellites include: the High Energy Transient Explorer HETE-2 (2000); ESA's INTErnational Gamma-Ray Astrophysics Laboratory INTEGRAL (2002); Swift (2004); and Fermi Gamma Ray Space Telescope (formerly Gamma-Ray Large Area Space Telescope GLAST, 2008). The Compton Gamma Ray Observatory CGRO (1991–2000) was a NASA mission, one of the Great Observatories in space, and included 4 instruments including Burst And Transient Source Experiment BATSE. (Graphs of bursts from this platform are shown at *IMA2* p. 543.) NASA's Gamma-ray Coordinates Network³¹³ keeps track of these events and provides rapid notification to subscribers.

These bursts are detected with current instruments about once per day, in energies ranging from 1 keV (soft X-rays) to many GeV. NASA describes two types: (1) **short-hard GRBs** (< 2 s, averaging 0.3 s)— associated with neutron star–neutron star mergers or neutron star–black hole mergers but not with SNe; and (2) **long-soft GRBs** (> 2 s, averaging 30 s, and having less energy in gamma rays than short-hard)—sometimes associated with SNe or even **hypernovae** or **collapsars**.³¹⁴ Overall, GRBs last over a range of 0.01 s to 1000 s. The integrated energy flux per unit area of detector is expressed as **fluence**, and has measured as great as 10^{-7} Jy m⁻².

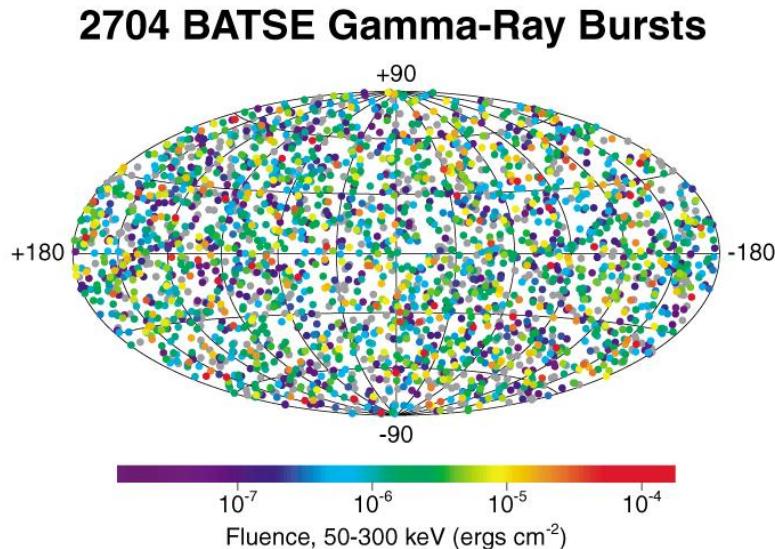
³¹² Klebesadel RW et al, “Observations of Gamma-Ray Bursts of Cosmic Origin”, *Astrophysical Journal*, vol. 182, p.L85, DOI: 10.1086/181225.

See also http://en.wikipedia.org/wiki/Vela_%28satellite%29

³¹³ <http://gcn.gsfc.nasa.gov/about.html>

³¹⁴ http://imagine.gsfc.nasa.gov/docs/science/know_11/grbs_explosion.html

GRBs have been found to have an isotropic angular distribution across the celestial sphere, and therefore must arise extragalactically (see graphs of all-sky distribution of GRBs and output of GRB 970228 below). Because of their great distances yet strong signal, GRBs are among the most energetic events in the universe, comparable in energy to core-collapse SNe (*IMA2* p. 547).



All-Sky Map Showing Isotropic Distribution of 2704 GRBs detected by BATSE,
a part of CGRO, over a 9 year period (NASA, 2007)³¹⁵

Some long-soft GRBs are caused by SNe—for example, GRB 980425 in 1998 was shown to be associated with a distant unusually strong SN, SN 1998bw,³¹⁶ thought to be a Type Ib or Ic (*IMA2* p. 547 and *here*³¹⁷). However, although the GRB was atypically weak, the SN was unusually strong for a Type Ib/c, estimated at 2 to 6×10^{45} J. Another supernova associated with a GRB (GRB 030329) was SN 2003dh, although these correlations appear to be rare. Most long-duration GRBs are not associated with observable SNe. “Yet the indirect evidence ... supports the consensus view that most long-duration GRBs arise from the death of massive stars.”³¹⁸

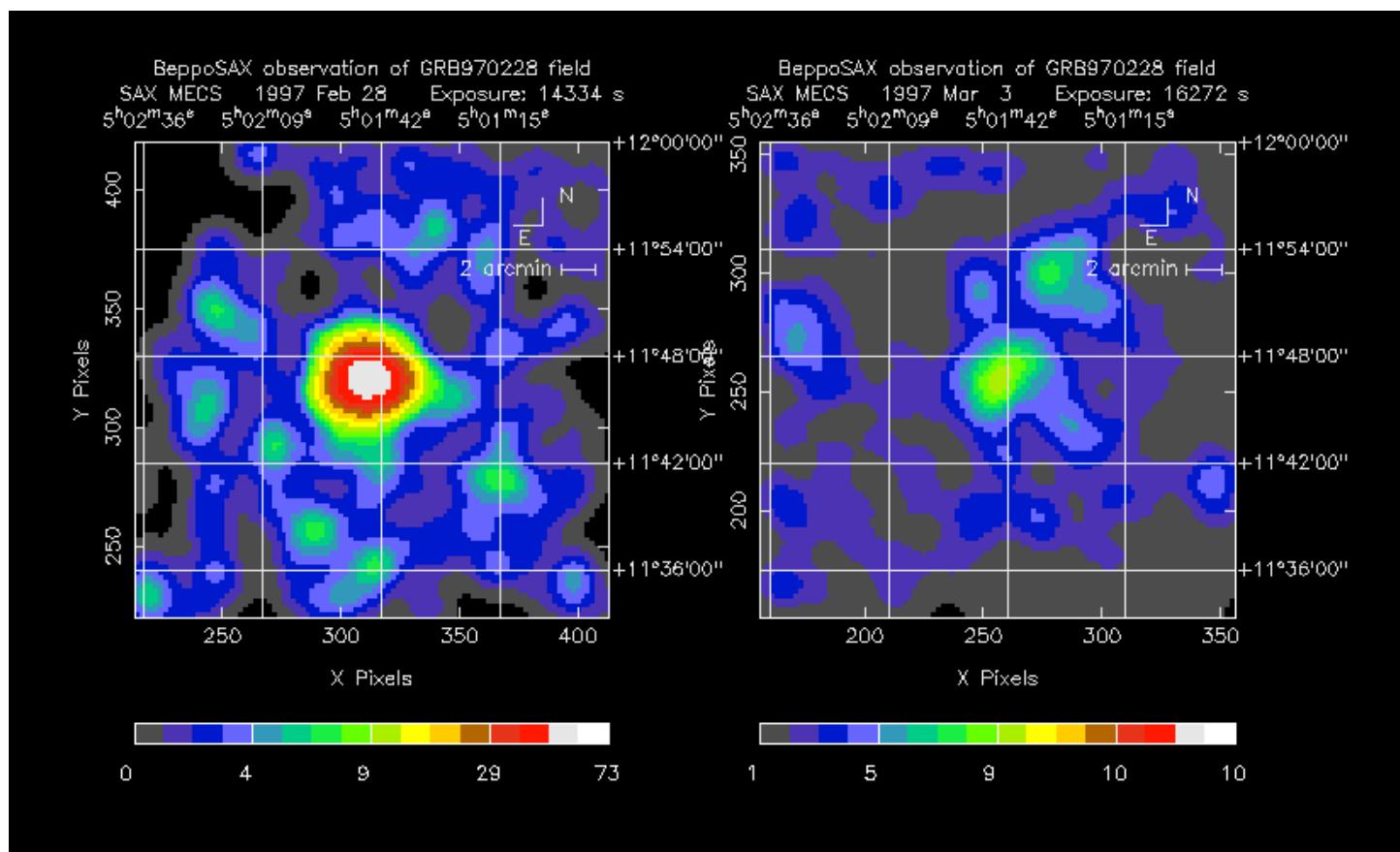
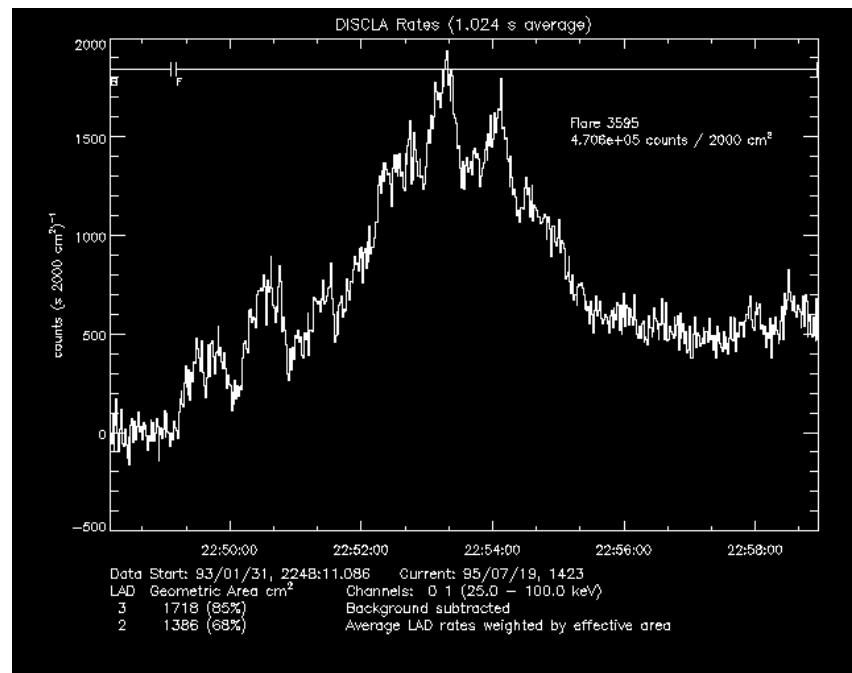
³¹⁵ <http://heasarc.gsfc.nasa.gov/docs/cgro/batse/>

³¹⁶ <http://www.ing.iac.es/PR/newsletter/news2/gamma.html>

³¹⁷ <http://www.nature.com/nature/journal/v395/n6703/full/395670a0.html>

³¹⁸ Woosley SE and Bloom JS, “The Supernova–Gamma-Ray Burst Connection”, *Annual Review of Astronomy and Astrophysics*, Vol. 44: 507–556, 2006, DOI: 10.1146/annurev.astro.43.072103.150558

The graph³¹⁹ to the right illustrates the gamma ray output for a strong GRB (CGRO Flare #3595) occurring on 31 January 1993. It had a peak count rate of 1956 counts s⁻¹ over 2000 cm² in the range 25 to 100 keV and lasted about 6 min:



Rapidly decaying X-ray emissions related to the 1997 GRB 970228³²⁰
(not the same GRB as graphed just above).

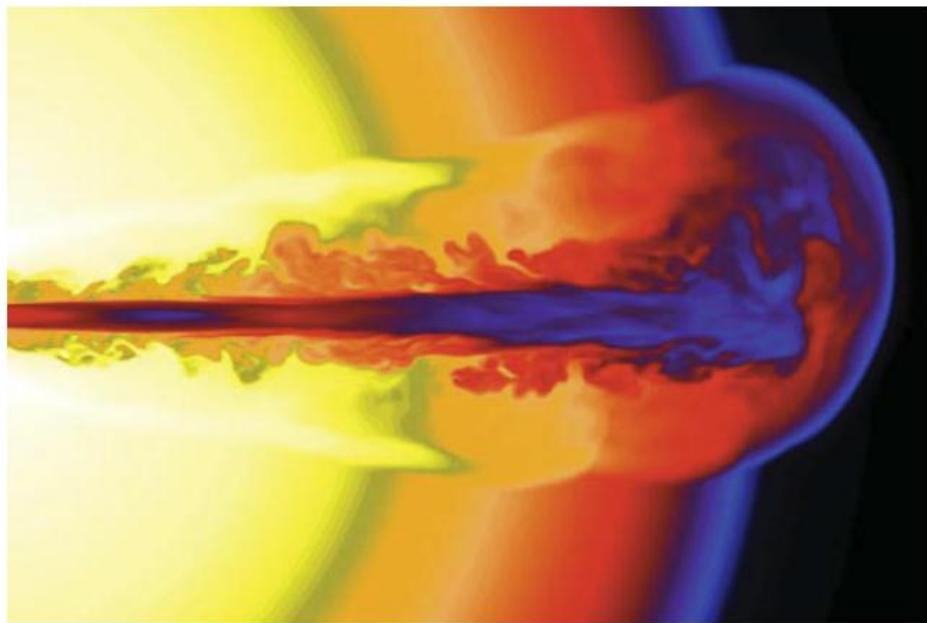
³¹⁹ ftp://umbra.nascom.nasa.gov/pub/batse/batsegifs/batse_3595.gif

³²⁰ <http://www.mpe.mpg.de/~jcg/grb970228.html>

The GRB of 2/28/1997, which lasted < 135 s as GRB 970228, was detected by the BeppoSAX satellite (in orbit 1996–2003), initially as a gamma ray burst. Subsequent X-ray detection (as 1SAX J0501.7+1146) was done 8 hours later. The MECS instrument detects 2–10 keV (soft X-rays). Correlation with optical counterparts localized the burst to a distant galaxy (though no SN was observed). The first image above³²¹ is of X-rays at 8 hour after the GRB and shows the pattern of marked increased counts in a region localized to the constellation of Orion. The second image is 2 days later, and shows a marked fall-off in X-ray count rate.

Mechanism of Long-Soft GRBs

The Long-Soft GRBs are thought to result from relativistically beamed jets (previously discussed). Highly directional jets reduce the amount of total energy that must be emitted to produce the GRB observed. For a Lorentz factor of $\gamma = 100$ (velocity= 0.99995c) and assuming we are in the relativistically beamed jet, the required energy output would be reduced by $1/\gamma^2$ or $1/10,000$ (*IMA2* p. 548). According to the hypernova or collapsar model of Woosley and Bloom (2006), a sufficiently massive star (such as a WR) undergoing a SN may leave a black hole with a collimating debris-laden disk and a magnetic field which together produces a highly relativistic beamed jet. This emerging jet interacts with overlying infalling envelope material, producing bursts of gamma rays. (Some details omitted.) The following image depicts a model of a jet emerging from a black hole in the center of a $15 M_{\odot}$ WR star undergoing a supernova (Woosley and Bloom, 2006):



Break out of a relativistic γ -ray burst jet with energy 3×10^{50} erg s^{-1} 8 s after it is launched from the center of a $15 M_{\odot}$ WR star. The radius of the star is 8.9×10^{10} cm and the core jet, at infinity, will have a Lorentz factor $\Gamma \sim 200$. Note the cocoon of mildly relativistic material that surrounds the jet and expands to larger angles. Once it has expanded and converted its internal energy this cocoon material will have Lorentz factor $\Gamma \sim 15$ –30. An off-axis observer may see a softer display dominated by this cocoon ejecta. If the star were larger or the jet stayed on a shorter time, the relativistic core would not emerge, though there would still be a very energetic, highly asymmetric explosion. (Zhang, Woosley & Heger 2004.)

³²¹ Gamma-ray burst 970228:

- <http://www.asdc.asi.it/bepposax/first/grb970228.html> including X-ray image shown
- http://www.sron.nl/divisions/heasax/htdocs/science_grb970228.html describes the GRB

Cosmic Rays and Solar Energetic Particles

Cosmic rays were discovered by Austrian physicist Victor F. Hess (1883–1964) in 1912, using a precision electroscope³²² carried aloft in a high-altitude balloon. “He showed that the level of radiation decreased up to an altitude of about 1 km, but above that the level increased considerably, with the radiation detected at 5 km about twice that at sea level. His conclusion was that there was radiation penetrating the atmosphere from outer space, and his discovery was confirmed by Robert Andrews Millikan in 1925, who gave the radiation the name ‘cosmic rays’.”³²³

The term *Cosmic ray* is a misnomer because primary cosmic rays initially consist of single charged particles (mostly protons), not rays of radiation (photons). Most are protons, a much smaller amounts are alphas, and most of the rest are nuclei of higher Z including Fe, Ni, and others up to uranium U. (See graph below for more precise relative proportions of various nuclei over a range of energies.³²⁴) Small numbers of high energy electrons, positrons, muons, pions, and lambda baryons, possibly antiprotons, etc. also qualify as primary cosmic rays—they are more commonly produced as secondary cosmic rays. Cosmic rays are variously defined, but the following may be suitable: “The cosmic radiation incident at the top of the terrestrial atmosphere includes all stable charged particles and nuclei with lifetimes of order 10^6 years or longer. Technically, ‘primary’ cosmic rays are those [stable charged] particles accelerated at astrophysical sources and ‘secondaries’ are those particles produced in interaction of the primaries with interstellar gas.”³²⁵ The defining lower energy boundary (if any) for primary cosmic rays is unclear to me, but appears to be about 10^7 eV (0.01 GeV) or 10^8 eV (0.1 GeV). In addition, there seems to be some uncertainty as to whether very high energy gamma rays can ever be considered as primary cosmic rays (probably not).

Primary cosmic rays lead to a shower or cascade of secondary cosmic ray particles, neutrinos, and photons, as a result of collisions typically in the upper atmosphere.

A single cosmic ray particle can have energy in excess of 10^{20} eV (16 J)—these are called ultra-high-energy cosmic rays (UHECR). One cosmic ray particle, the “Oh-My-God Particle” detected by ground-based Fly’s Eye II,³²⁶ was probably a proton. It had energy of 3.2×10^{20} eV or 51 J, equivalent to the KE of a 142 g hardball baseball traveling at 96 km/h.) This particle was moving so fast that in its rest frame, the time for it to travel from the Virgo Cluster would be only 1.15 hours.³²⁷ Such extremely high energy cosmic rays are not deflected by the magnetic field. However, the energy of most cosmic rays is between 10^7 eV and 10^{10} eV (see graph).

Most primary cosmic rays arise outside the Solar System, but the Sun is the source of relatively low energy cosmic rays (“**Solar energetic particles SEP**” or “**solar cosmic rays**”) arising from solar flares, recombination events, and CMEs.³²⁸ They consist of protons (plus generally lower energy electrons), and have energies of a few tens of keV to rarely above 10 GeV (10^{10} eV), and many are excluded of diverted to the poles by the Earth’s magnetic field.³²⁹ Typical peak SEP energies are about 10^7 eV (IMA2 p. 551)

³²² <http://en.wikipedia.org/wiki/Electroscope>

³²³ http://en.wikipedia.org/wiki/Victor_Francis_Hess

³²⁴ <http://pdg.lbl.gov/2011/reviews/rpp2011-rev-cosmic-rays.pdf> an excellent overall review from the Particle Data Group by K. Nakamura et al, 2011

³²⁵ <http://pdg.lbl.gov/2011/reviews/rpp2011-rev-cosmic-rays.pdf>

³²⁶ <http://www.cosmic-ray.org/reading/flyseye.html>

³²⁷ <http://www.fourmilab.ch/documents/OhMyGodParticle/>

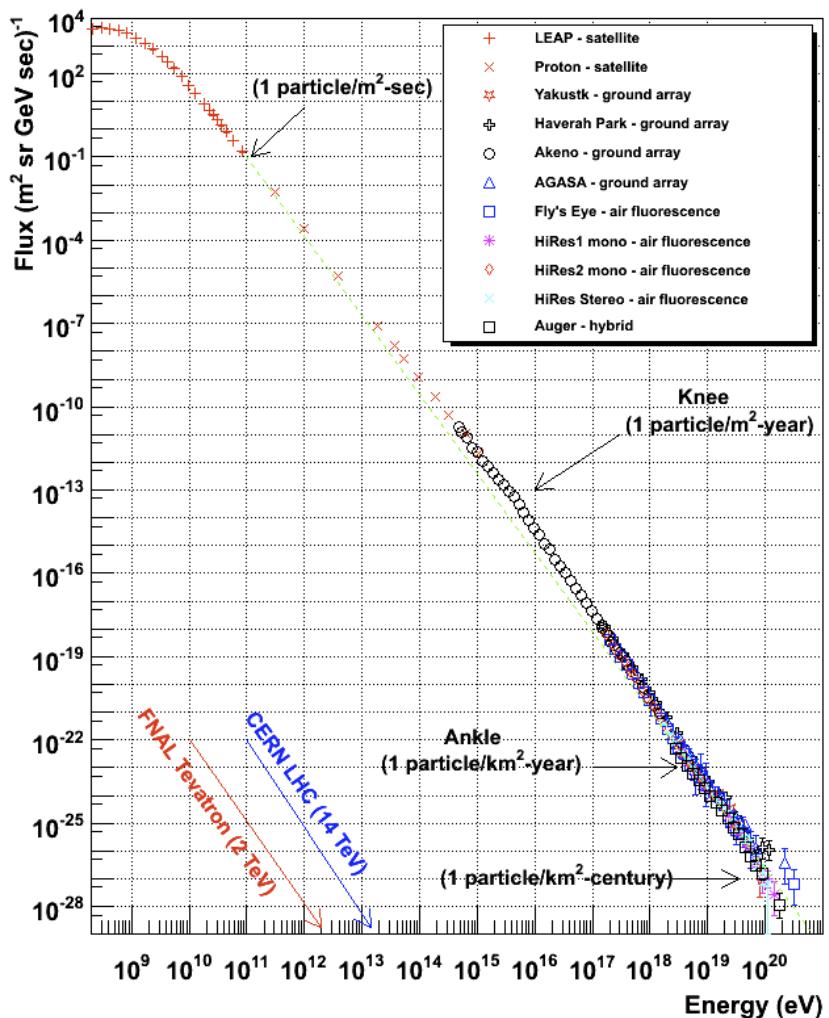
³²⁸ Solar Energetic Particles SEPs:

- <http://helios.gsfc.nasa.gov/sep.html>
- <http://solarphysics.livingreviews.org/open?pubNo=lrsp-2006-2&page=articlesu15.html>

³²⁹ <http://www-spof.gsfc.nasa.gov/Education/wsolpart.html>

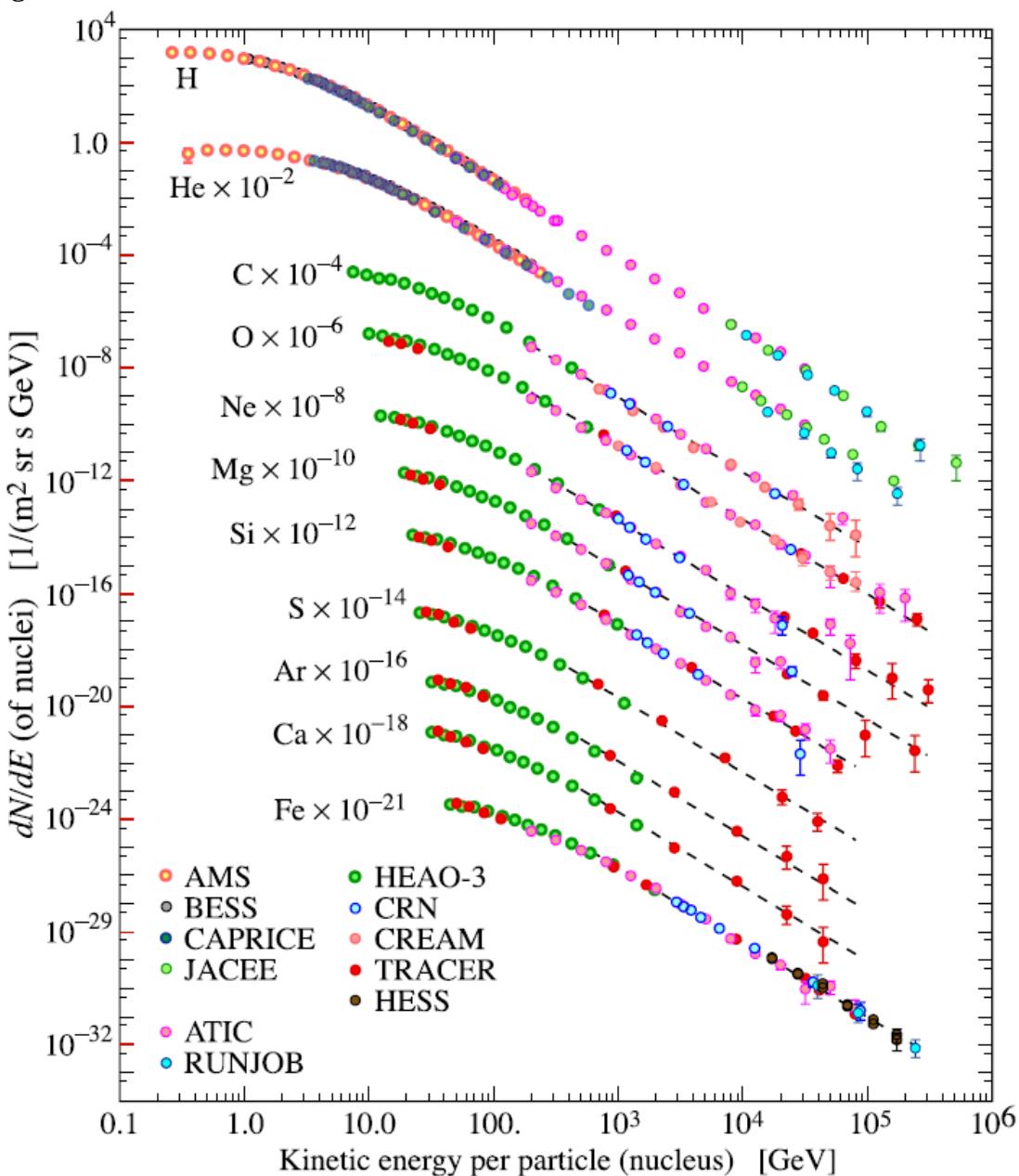
The graph to the right shows the steady decrease in the numbers of cosmic rays (expressed in particles per ($\text{m}^2 \text{ sr GeV sec}$)) with increasing kinetic energy, over the range of about 10^8 eV (1.6×10^{-10} J) to the current record-holder of 3.2×10^{20} eV (51 J)).³³⁰ The far left of the graph includes the upper range of Solar Energetic Particles. The knee region is associated with SNe. The rest energy of a proton is almost 10^9 eV, and is presumably not included in the particle energies shown in the graph.

Ultra-high energy cosmic rays are hypothesized to arise from Larmor gyrating particles repeatedly encountering the advancing shock wave front of a SN. Particles greater in energy than 10^{15} eV experiencing the Lorentz force arising from the typical magnetic field of the ISM are not likely to remain bound in the dimensions of a SN remnant—they are therefore candidates for becoming extragalactic cosmic rays (IMA2 p. 552). Other causative mechanisms may apply to cosmic rays with energies much above the knee at 10^{15} eV. For instance, at the ankle of the graph near 10^{19} eV, it is suggested that acceleration in the vicinity of neutron stars or black holes may be involved (IMA2 p. 553)



³³⁰ <http://www.physics.utah.edu/~whanlon/spectrum.html>, data apparently taken from Cronin J, Gaisser TK, Swordy SP. "Cosmic Rays at the Energy Frontier." *Scientific American* 276, 1, 44, 1997. Page 127 of 249

The following is a recent 2011 graph showing numbers of primary cosmic ray nuclei versus KE for multiple nuclear species. Note that energies for these nuclei begin above 10^8 eV and that protons ("H") predominate, followed by small amounts of He, and much lower amounts of C, O, and Ne. N is not included in this graph, of uncertain significance.³³¹



³³¹ <http://pdg.lbl.gov/2011/reviews/rpp2011-rev-cosmic-rays.pdf>

The Degenerate Remnants of Stars (Chapter 16)

White Dwarfs (WDs)

First Discovery

"The first white dwarf discovered was in the triple star system of 40 Eridani, which contains the relatively bright main sequence star 40 Eridani A, orbited at a distance by the closer binary system of the white dwarf 40 Eridani B and the main sequence red dwarf 40 Eridani C. The pair 40 Eridani B/C was discovered by William Herschel on 31 January 1783."³³²

The next to be discovered was Sirius B. After measuring the distance to 61 Cygni by stellar parallax, Friedrich Wilhelm Bessel in 1838 measured the distance to Sirius. It is the brightest star in the sky, and is actually a binary star system consisting of a white A1V MS star and a faint white dwarf at about 2.64 pc. Bessel detected a wobble of Sirius (actually Sirius A) and deduced in 1844 the presence of an unseen companion (Sirius B). The wobble is due to the elliptical orbit with period 50 y of Sirius A about the shared barycenter. Alvan Graham Clark in 1862 actually observed the predicted companion, which was near periastron. Current estimate of the Sirius B mass is $1.053 M_{\odot}$. The pair differ greatly in luminosity: $L_A = 23.5 L_{\odot}$ while $L_B = 0.03 L_{\odot}$. Spectroscopy in 1915 revealed that Sirius B is an unexpectedly hot blue-white star with $T_{eff} = 27,000$ K, a **white dwarf**. By Stefan-Boltzmann Law, the radius was estimated at only $0.008 R_{\odot}$, its average density estimated at an enormous $3 \times 10^9 \text{ kg m}^{-3}$, and its surface gravitational acceleration estimated at a very intense $4.6 \times 10^6 \text{ m s}^{-2}$. The spectrum from the surface atmosphere exhibits H absorption lines with pressure/collisional broadening characteristic of DA WD stars. (H is otherwise depleted) in these WDs, *IMA2* p. 557 and p. 270.

Classification of White Dwarfs

WDs occupy the lower right corner of the H-R diagram. Their spectral type is D for dwarf or degenerate. They are subdivided first by spectroscopic features, using a second letter:

- DA:** WDs having strong H Balmer absorption lines, typically pressure-broadened
- DB:** WDs having He I absorption lines, no H lines
- DO:** WDs having He II absorption lines, no H lines
- DC:** WDs having no absorption lines, only a featureless continuum
- DQ:** WDs having carbon features and spectral lines
- DZ:** WDs having metals features and spectral lines
- DX:** WDs insufficient or unclear data preventing classification

A hybrid WD might be designated DAB, meaning having H and He I absorption lines—also, DAO, DAZ, or DBZ may be employed.

An added V means variable. For instance, **DAV** are pulsating white dwarfs (aka ZZ Ceti).

A temperature index number may be added, such as **DA2** (as with Sirius B). The temperature index number is computed by dividing 50,400 K by the effective temperature.³³³

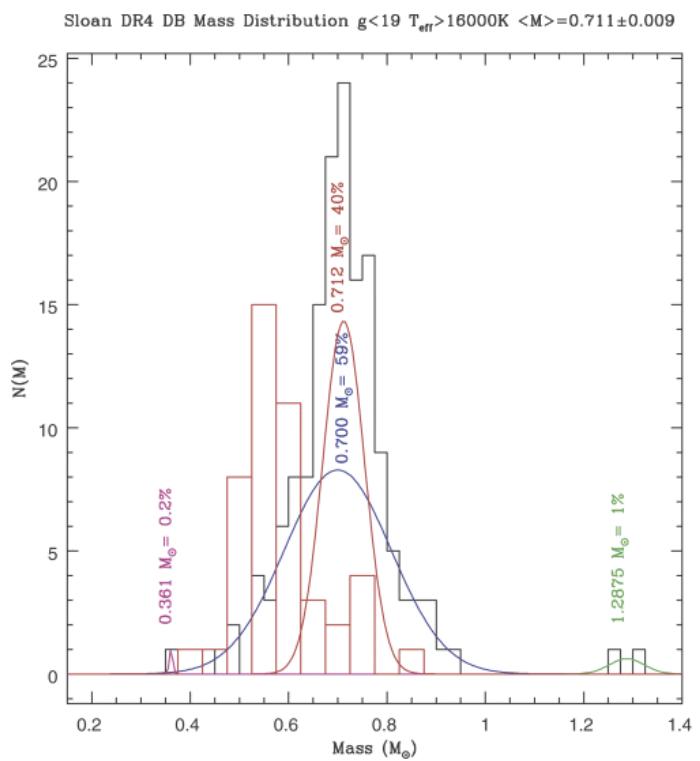
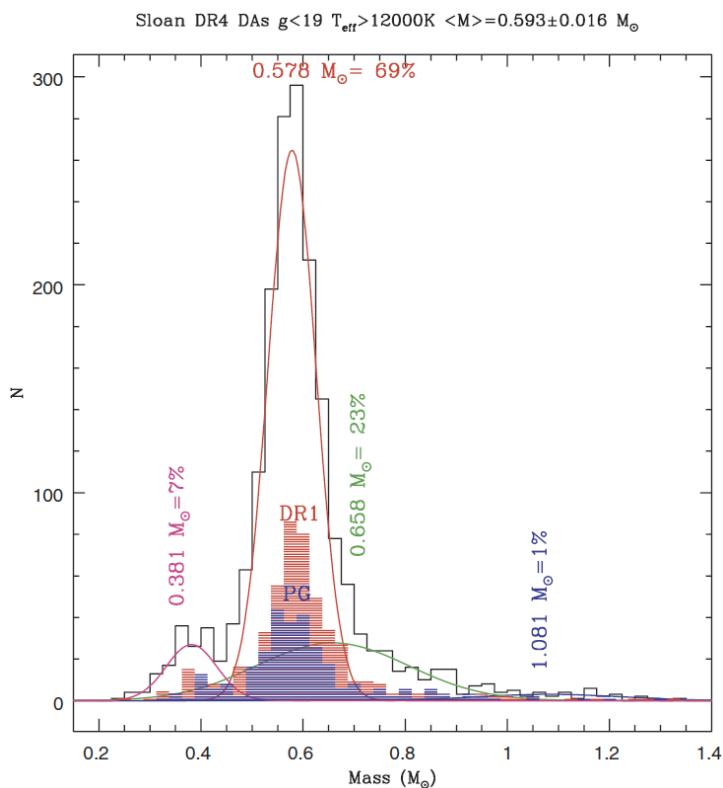
A luminosity class of VII is rarely applied.

The masses of DA WDs peak at a value of $0.56 M_{\odot}$ and 80% lie within 0.42 to $0.7 M_{\odot}$ (*IMA2* p. 560).

³³² http://en.wikipedia.org/wiki/White_dwarf

³³³ *ibid.*

The SDSS identified more than 7500 WDs that were included in the SDSS 2005 DR4 (see histograms to follow, all quotes and images in this table come from *this*³³⁴ article):



WD DAs: For the DAs, “The mean mass for the DA stars brighter than $g = 19$ and hotter than $T_{\text{eff}} = 12,000$ K is

$$\langle M_{\text{DA}} \rangle = 0.593 \pm 0.016 \text{ M}_{\odot}$$

In this graph, the main Gaussian fit peak is shown in red and smaller additional Gaussian fits are shown in other colors. 69% of DAs are enclosed by the red Gaussian curve, which extends from about 0.45 to about 0.7 M_{\odot} . (DR1 is SDSS data from earlier dataset published in 2004, PG is from Palomar Green survey in 2005.)

WD DBs: For the DBs with g [green band apparent magnitude] < 19 and $T_{\text{eff}} > 16000$ K, the average mass was found to be

$$\langle M_{\text{DB}} \rangle = 0.711 \pm 0.009 \text{ M}_{\odot}$$

“It appears the mean mass for DB white dwarf stars may be significantly larger than that for DAs. We also report the highest mass white dwarf stars ever found, up to 1.33 M_{\odot} .” The histograms are actual data, the 4 Gaussian fits approximate portions of their distributions. No prior studies included.

Central Conditions and Origin

The core or central pressure of Sirius B P_c is crudely estimated to be about $3.8 \times 10^{22} \text{ N m}^{-2}$ (IMA2 p. 559), which is more than a million times that of the Sun’s core P . The core temperature T_c of Sirius B is crudely estimated at around 7.6×10^7 K. Despite the high temperatures, it is deduced that thermonuclear reactions are not contributing to WD luminosity, and that their centers are incapable of fusion reactions at the estimated T_c and P_c . The core and much of remainder of most WDs (other than the atmosphere) are thought to be primarily ionized C and O nuclei.

WDs arise from intermediate mass progenitor stars of less than 8 or 9 M_{\odot} , forming in the stellar cores in the AGB or post-AGB or after a planetary nebula stage that by shedding outer layers exposes the core.

³³⁴ Kepler SO, et al, “White dwarf mass distribution in the SDSS”, *Monthly Notices of the Royal Astronomical Society*, Volume 375, Issue 4, published online: 12 FEB 2007, DOI: 10.1111/j.1365-2966.2006.11388.x

Small WDs may retain unburned He and become **helium WDs**. Certain larger WDs of 8 to 10 M_⊕ may have fused C but not Ne, and consist largely of oxygen, neon, magnesium, and other intermediate mass elements. These are called **oxygen-neon-magnesium WDs**, and are detected by **ONeMg novae**). (IMA2 p. 560 and *here*³³⁵)

Surface Conditions and Spectra

As mentioned above, the spectral characteristics are determined by the composition of the atmospheres. Because of the extreme gravity, the H if any is present rises to the very top and heavier elements sink toward the core stratified by their density. This means the outer layer of H, which gives the DA its H spectrum, may be very thin and overly a thin layer of He, which overlies the O-C core. This stratification takes place quickly over about 100 years. Convection may mix He into the H atmosphere, diluting out the H and changing the type from DA to DB.

Pulsating WDs (DAV=ZZ Ceti variables, DBV=V777 Her variables, DOV)

DAV WDs (ZZ Ceti variables) have Teff of 10,000 to 12,000K, and periods of 100 to 1000 sec (as previously discussed). They pulsate in non-radial g-modes that affect temperature, and that affect the mean intensity of light emitted.³³⁶ The almost entirely horizontal displacement of g-modes is compatible with the intense gravitational pull, and little radial change occurs from pulsation. It is the hydrogen partial ionization zone that drives the pulsations of DAVs. For the hotter and more luminous DBVs (V777 Her variables), the helium partial ionization zone is responsible). The even hotter DOVs are found in the H-R diagram at the transition from the very hot PNNV stage (planetary nebula nuclei variable) to WDs, IMA2 p. 562 and *here*³³⁷) These stars oscillate in multiple modes simultaneously, a topic of great current research interest. The variation in light output is small.

The Physics of Degenerate Matter

Normal thermal gas pressure—given for an ideal gas by NkT/V, where N=number of particles—and radiation pressure—given by (4σT⁴/3c)—are entirely inadequate to support a WD against complete collapse. The mechanism that does provide the support is **electron degeneracy**, which was deduced by Ralph H. Fowler (1889–1944) in 1926, working with Paul Dirac (only one year after Wolfgang Pauli published his paper on the Pauli Exclusion Principle). The Exclusion Principle requires that the total wave function for two identical fermions is anti-symmetric with respect to exchange of the particles, or in simpler terms, that no two fermions can occupy exactly the same quantum state. For electrons in atoms, there can be no two with the same n, l, m_l, and m_s. For any 2 electrons attempting to have the same quantum numbers, the probability amplitude resulting from adding the two parts of the anti-symmetric wave function is simply zero.³³⁸

For fermions like electrons in a confined region such as a box, the wave functions are de Broglie standing waves described by three quantum numbers, the components of momentum in 3 directions. Electrons can have spin quantum number m_s = ±1/2, a fourth quantum number. No two electrons can have exactly the same 4 quantum numbers. As the temperature and therefore the thermal pressure decline, electrons fill up the lower unoccupied quantum states. A **completely degenerate** gas is one in which all of the lower energy states are occupied and none of the higher energy states are occupied—this occurs only at 0 K and can be expressed by ε ≤ ε_F. (IMA2 p. 563) For higher temperatures, some of the states with energy ε ≤ ε_F will be vacant, and some electrons will occupy states with energy ε > ε_F. Here, ε_F is the **Fermi Energy**, the maximum

³³⁵ http://en.wikipedia.org/wiki/White_dwarf

³³⁶ Winget DE, “Asteroseismology of white dwarf stars”, *Journal of Physics: Condensed Matter* Volume 10 Number 49 1998, doi:10.1088/0953-8984/10/49/014

³³⁷ <http://adsabs.harvard.edu/abs/2007arXiv0705.2115S> , Solheim et al, 2007

³³⁸ Pauli Exclusion Principle:

- http://en.wikipedia.org/wiki/Pauli_exclusion_principle
- <http://hyperphysics.phy-astr.gsu.edu/hbase/pauli.html>

electron energy found in a completely degenerate gas at 0 K. For higher temperatures, ϵ_F is the energy at which 50% of the states for that energy are occupied by electrons.

By applying the relation between momentum and de Broglie wavelength, $p=h/\lambda$, and the kinetic energy relationship to momentum, $\epsilon = p^2/2m$, a formula for Fermi energy is derived:

$$\epsilon_F = \frac{\hbar^2}{2m} (3\pi^2 n)^{2/3}$$

where m is the electron (or other Fermion) mass, and n is the number of electrons per unit volume. (IMA2 p. 565)

Despite the high temperatures of a WD, it is a good approximation at such high densities to assume complete degeneracy for the calculation, so that all but the most energetic electrons have $\epsilon \leq \epsilon_F$...

The **condition for degeneracy** becomes

$$\frac{T}{\rho^2} < \mathfrak{D} = 1261 \text{ K m}^2 \text{ kg}^{-2/3}$$

(assuming Z/A, the ratio of protons to nucleons for the gas, is 0.5). The smaller the LHS is compared to \mathfrak{D} (a quantity derived from constants and the composition of the gas), the more degenerate the gas is.

Applying this formula to standard models of the Sun's core, the LHS is 5500. The Sun has very low electron degeneracy, supplying only "a few tenths of a percent" of the central pressure (IMA2 p. 566). However, as it evolves along the SGB it will acquire much higher degeneracy, which peaks at the He core flash and the TP-AGB. (IMA2 p. 567)

In contrast, applying this formula to values for Sirius B's core, the LHS is only 37, much less than 1261 in the same units, and complete degeneracy is a valid assumption.

As electrons are forced into more tightly defined positions, their momenta increase (by Heisenberg Uncertainty). The rising momenta contribute a temperature-independent pressure, the **degeneracy pressure**, for which the derived final formula is (IMA2 p. 569)

$$P_{\text{deg}} = \frac{(3\pi^2)^{2/3} \hbar^2}{5} \frac{1}{m_e} \left[\left(\frac{Z}{A} \right) \frac{\rho}{m_H} \right]^{5/3}$$

where m_e is mass of an electron, m_H is mass of hydrogen (almost the same as mass of a proton), Z is average number of protons per nucleus, A is average number of nucleons per nucleus. It is this degeneracy pressure—to which is added a much smaller thermal gas pressure and radiation pressure—which maintains hydrostatic equilibrium against gravity in a white dwarf.

For the extreme relativistic limit, where electron speeds are close to c, IMA2 p. 570 gives an alternate formula for degeneracy pressure:

$$P_{\text{deg}} = \frac{(3\pi^2)^{1/3}}{4} \hbar c \left[\left(\frac{Z}{A} \right) \frac{\rho}{m_H} \right]^{4/3}$$

Comparing the Sun with the WD Sirius B, I calculate (using formulas from IMA2, p. 569, 291, and 295 and assumed values shown) the following (all units are MKS):

	Result N m⁻²	Z / A	Mean MW μ incl. electrons	Core Density ρ
Sun Core Gas Pressure	2.74E+16		? 0.723 ³³⁹ uncertain estimate	152700
Sun Core Radiation P	1.53E+13			
Sun Degeneracy P	2.17E+15	0.5		

³³⁹ <http://iopscience.iop.org/0004-637X/670/1/872/fulltext/72165.text.html>
Page 132 of 249

	Result N m⁻²	Z / A	Mean MW μ incl. electrons	Average Density ρ
Sirius B Core Gas Pressure	4.78E+20		? 1.6 [14/9] uncertain estimate	3.00E+09
Sirius B Core Rad. P	2.04E+14			
Sirius B Degeneracy P	3.11E+22	0.5		

If these rough calculations are in the right ball park (and I am somewhat unclear about the average molecular weight term), the Sun is shown to have high gas pressure, negligible radiation pressure compared to its gas pressure, and small degeneracy pressure compared to its gas pressure. In contrast, Sirius B has very high degeneracy pressure (much higher than the Sun's core gas pressure), modest gas pressure compared to its degeneracy pressure (though much higher than the Sun's gas pressure), and negligible radiation pressure compared to degeneracy gas pressure (but still higher than the Sun's radiation pressure).

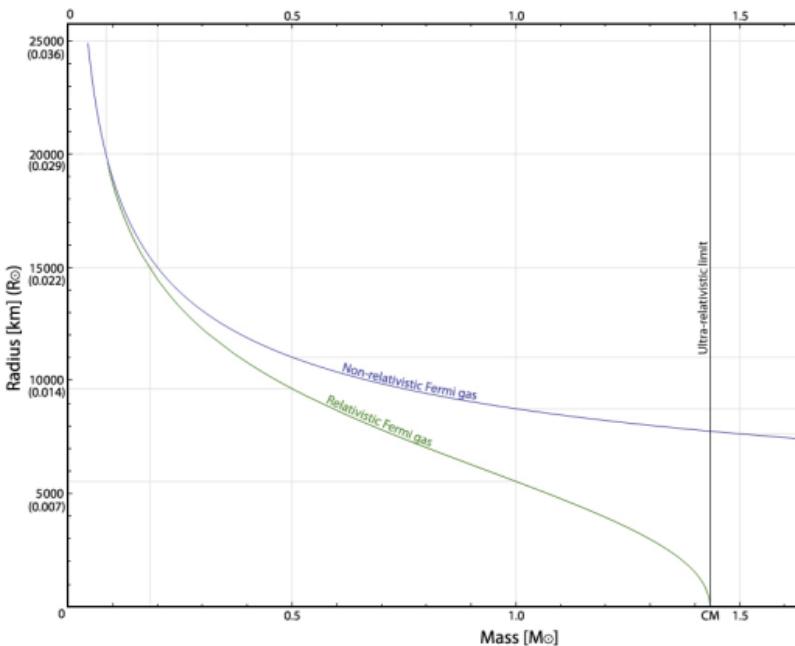
The Chandrasekhar Limit for White Dwarfs

Subrahmanyan Chandrasekhar (1910–1995) calculated in 1930 an upper size limit for a white dwarf, beyond which the core is not adequately supported by its electron degeneracy pressure, resulting in a collapse to a much denser neutron star.

$$R_{WD} \approx \frac{(18\pi)^{2/3}}{10} \frac{\hbar^2}{Gm_e M_{WD}^{1/3}} \left[\left(\frac{Z}{A} \right) \frac{1}{m_H} \right]^{5/3}$$

The radius R_{wd} of a white dwarf is thus inversely proportional to $M_{wd}^{1/3}$, and in the non-relativistic approximation, $M_{wd} * V_{wd} = \text{constant}$. Thus, volume of a white dwarf is approximately inversely proportional to its mass M . For large enough mass, relativistic considerations apply, dynamical instability results (i.e., there is no stable solution), the volume shrinks drastically, and the white dwarf collapses to become a neutron star. (This is related to the collapse of a depleted stellar core lacking adequate total core pressure, causing a core-collapse supernova.) *IMA2* p. 571 gives an approximate formula for this M_{Ch} , but its underestimates the currently accepted Chandrasekhar Limit of $M_{ch} = 1.44 M_{\odot}$.

The graph to the right³⁴⁰ show the relationship of radius to mass of white dwarfs. The lower curve depicts the relativistic case, for which mass falls to near-zero radius at $M = M_{Ch}$. This is similar to the graph at *IMA2* p. 571. I have not determined the significance of the “ultra-relativistic limit”.



The pressure P in a degenerate core is independent of T , and as noted earlier, a runaway thermonuclear reactions can transiently occur at the “helium core flash” of an evolving star at the Red Giant tip of the RGB phase.

³⁴⁰ http://en.wikipedia.org/wiki/Chandrasekhar_limit

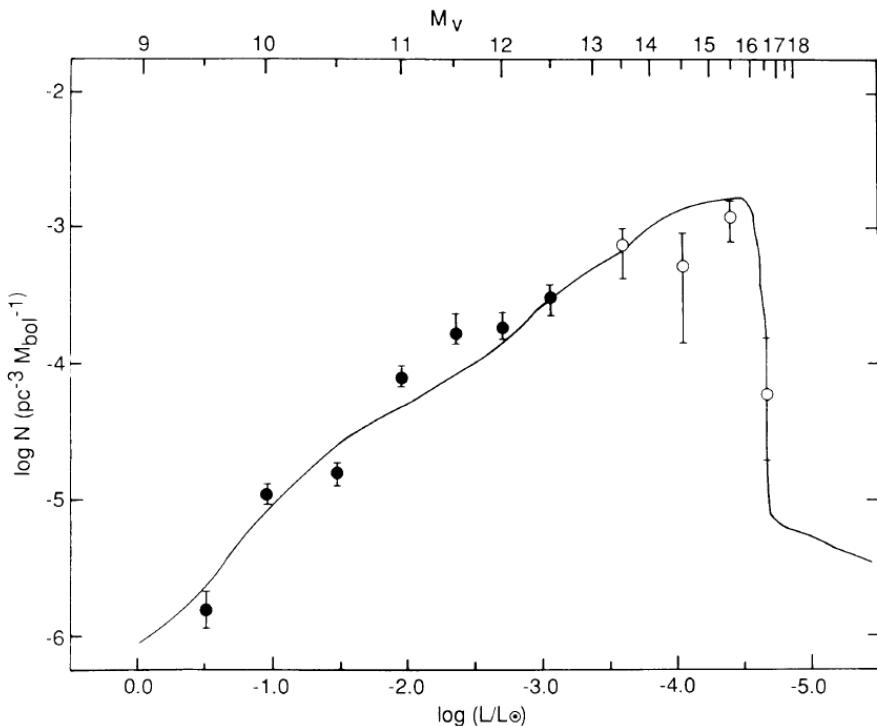
Cooling of White Dwarfs

In a WD, degenerate electrons can travel long distances (because most of the electron states along their way are already occupied). Electron conduction rather than photon radiation therefore is the principle means of transporting heat toward the surface in the highly degenerate core, and its efficiency causes the interior of a WD to be nearly isothermal. The outer thin non-degenerate envelope of the WD is the location where T suddenly begins to fall and the T gradient leads to superficial convection. “The surface of a WD cools more slowly than its isothermal interior as the star’s thermal energy leaks into space.” (IMA2 p. 574, not absolutely clear to me)

Most of the energy that is given up from cooling is in the nuclei, as the electrons are already mostly in their lowest energy states and can go no lower. A rough calculation of available energy to be radiated away is 6×10^{40} J, and this would be lost by luminosity at a cooling timescale of more than about 170 Ma. A cooling curve for a $0.6 M_{\odot}$ WD (IMA2 p. 576) shows $\log(L/L_{\odot})$ dropping from 1 initially to about -4 (1/10,000) in about 5 Ga and to -5 (1/100,000) in about 15 Ga.

A fascinating aspect of WD cooling is that it will eventually undergo crystallization from the inside out. This is the C-O core, which forms body-centered cubic lattice (BCC) crystals. The release of latent heat of crystallization temporarily slows down the cooling curve at the “knee” of the curve... One of the more memorable quotes in the textbook states, “Thus the ultimate monument to the lives of most stars will be a ‘diamond in the sky,’ a cold, dark, Earth-size sphere of crystallized carbon and oxygen floating through the depths of space.” (IMA2 p. 576) Incidentally, the packing factor for a BCC crystal is 0.680,³⁴¹ tighter packing than that seen with terrestrial diamonds (which have a diamond cubic crystal structure with packing factor of only ~0.34).³⁴²

Cooling of a pulsating WD can be observed in the form of slowing of pulsations. It is inferred from the number density of WDs of various ages (graph to right, vertical axis)³⁴³ that the number of WDs falls off rapidly at $\log(L/L_{\odot}) = -4.5$. This leads to a conclusion that WDs in the MW galactic disk first began to form about 9 Ga ago, and that their progenitor disk stars began to form at about 9.3 Ga ago (based on a mean pre-white-dwarf lifetime of 0.3 Gyr). (However, MW globular cluster stars have been shown to be up to 3 Ga older.) (IMA2 p. 578)



³⁴¹ http://en.wikipedia.org/wiki/Cubic_crystal_system

³⁴² http://en.wikipedia.org/wiki/Diamond_cubic

³⁴³ Winget DE, et al, “An independent method for determining the age of the universe”, *Astrophysical Journal, Part 2 - Letters* vol. 315, April 15, 1987, p. L77-L81. <http://adsabs.harvard.edu/doi/10.1086/184864>

Neutron Stars (NS)

The neutron was discovered by James Chadwick (1891–1974) in 1932 and announced in a letter to the editor.³⁴⁴ Neutron stars were first proposed in 1934 by Walter Baade (1893–1960) and Fritz Zwicky (1898–1974), also as a letter to the editor.³⁴⁵ They suggested that “super-novae” represent a transition from ordinary stars to neutron stars.

Neutron stars are stellar remnants that form when increasingly degenerate stellar cores approach a mass near or exceeding the Chandrasekhar Limit ($\sim 1.4 M_{\odot}$). The core collapses, causing a core-collapse supernova (Type Ib, Ic, or II, see Chapter 15 discussion) and leaves a neutron star stellar remnant. They resemble a single confluent nucleus consisting mainly of neutrons, about 10^{57} neutrons total, held together by gravity which is opposed by neutron degeneracy pressure. Neutrons are also fermions subject to Pauli Exclusion principle.

The radius of a neutron star is given by

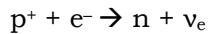
$$R_{WD} \cong \frac{(18\pi)^{2/3}}{10} \frac{\hbar^2}{Gm_e M_{NS}^{1/3}} \left[\frac{1}{m_H} \right]^{8/3}$$

This formula, which assumed a constant density, underestimates the radius somewhat. For a $1.4 M_{\odot}$ NS, the formula gives a radius of 4.4 km, but the best estimates are 10 to 15 km.³⁴⁶ For a NS of radius 10 km and mass of $1.4 M_{\odot}$, the average density would be $6.65 \times 10^{17} \text{ kg m}^{-3}$. (At the center, the density is estimated at $10^{18} \text{ kg m}^{-3}$, see below.) By comparison, an atomic nucleus has a significantly smaller density of about $2.3 \times 10^{17} \text{ kg m}^{-3}$ (IMA2 p. 578 and [here](#)³⁴⁷), and the density of a nucleon is estimated at $3.0 \times 10^{17} \text{ kg m}^{-3}$ (assuming a neutron has a classical radius of 1.1 fm and a mass of $1.67 \times 10^{17} \text{ kg}$). Therefore, the neutrons are tightly packed together, perhaps even overlapping.

The gravitational force at the surface has typical values of a few $\times 10^{12} \text{ m s}^{-2}$ up to a maximum of $7 \times 10^{12} \text{ m s}^{-2}$ (which is more than 10^{11} times of that of Earth, 9.78 m s^{-2}).³⁴⁸ An object dropped from 1 m would accelerate to an impact velocity of $1.9 \times 10^6 \text{ m s}^{-1}$. Calculations should use general relativity for most accurate results. See Chapter 18 summary (and IMA2 p. 661) for estimation of mass release by matter infalling to a NS.

The Equation of State

The following discusses the sequence whereby an “iron white dwarf” at the center of a massive supergiant star progresses to a neutron star. At the high pressure and density $\approx 10^9 \text{ kg m}^{-3}$, such as seen in such an iron core (IMA2 p. 579), the electrons become relativistic and interact with protons (“electron capture”) to create neutrons:



The relativistic electrons supply the energy needed for the greater mass of the products, which differs from that of the reactants by about 0.78 MeV. The density at which this reaction begins to occur is thought to be $1.2 \times 10^{10} \text{ kg m}^{-3}$ (IMA2 p. 580). Note that this value is somewhat higher than the estimated average white dwarf density. (I have not yet found good estimates of core white dwarf density.) A more detailed calculation determines that the actual density needed in an iron core for ^{56}Fe protons to capture relativistic free electrons is even higher, $> 10^{12} \text{ kg m}^{-3}$. This process of **neutronization** reduces the Coulomb repulsion as a result of declining proton packing density, and leads to nucleosynthesis of a series of nuclides: ^{56}Fe , ^{62}Ni , ^{64}Ni , ^{66}Ni , ^{86}Kr , ..., ^{118}Kr . (IMA2 p. 580 and [here](#)³⁴⁹). The neutron-rich nuclei do not decay by beta-decay, because there

³⁴⁴ Chadwick J, “Possible Existence of a Neutron”, *Nature* 129:312, 1932. doi:10.1038/129312a0

³⁴⁵ Baade W and Zwicky F, “Remarks on Super-Novae and Cosmic Rays”, *Phys. Rev.* 46, 76–77 (1934)

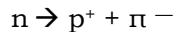
³⁴⁶ See discussion of radiation radius given in Lattimer JM and Prakash M, “Neutron Star Structure and the Equation of State”, *The Astrophysical Journal*, 550:426–442, March 20, 2001, <http://arxiv.org/abs/astro-ph/0002232>

³⁴⁷ Nuclear Density: for $^{12}\text{Carbon}$ = $(1.99264824 \times 10^{-26}) / ((4/3) * \pi * ((1.2 * (12^{(1/3)}) * 10^{-15}))^3) = 2.29 \times 10^{17} \text{ kg m}^{-3}$, see <http://www.chemistryexplained.com/Ar-Bo/Atomic-Nucleus.html>

³⁴⁸ IMA2 p. 578 and http://en.wikipedia.org/wiki/Neutron_star

³⁴⁹ Wallerstein G et al, “Synthesis of the elements in stars: forty years of progress”, *Rev. Mod. Phys.* 69, 995–1084 (1997).

are no vacant electron states available to take the electron. At $4 \times 10^{14} \text{ kg m}^{-3}$, the core is a lattice mixture of neutron-rich nuclei, free neutrons (**neutron drip**), and free degenerate electrons. The free neutrons pair up to form bosons which are not subject to Pauli Exclusion Principle and therefore can all crowd into a **superfluid** state of lowest energy. At a density of $4 \times 10^{15} \text{ kg m}^{-3}$, **neutron degeneracy pressure** exceeds electron degeneracy pressure. By density $\rho_{\text{nuc}} \approx 2 \times 10^{17} \text{ kg m}^{-3}$, there is no longer a meaningful distinction between inside versus outside nuclei, and nuclei dissolve. The charged protons constitute a **superconducting** fluid. The core has therefore become a fluid mixture of superfluid free neutrons, superfluid superconducting free protons, and relativistic free electrons, in number ratios of 8:1:1. Pions begin to form when density is more than twice nuclear density, as a result of decay of neutrons:



Modeling becomes increasingly difficult.

A “recent calculation” for a $1.4 M_\odot$ neutron star with radius 10.6 km (*IMA2* p. 582-3, no reference given, see also [here](#)³⁵⁰) suggests that there are the following layers:

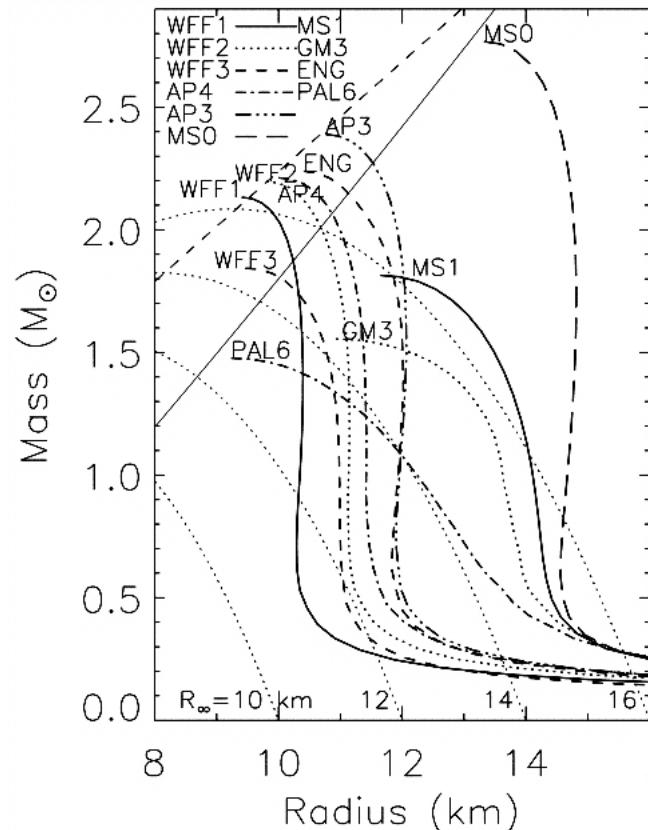
- **Outer Crust**, 10.3 to 10.6 km:
Consists of heavy nuclei (ions), nonrelativistic and relativistic degenerate electrons, a liquid or solid lattice form, mostly ^{56}Fe near the surface, increasing neutron-rich nuclei with depth. The bottom is the site of neutron drip at $\rho = 4 \times 10^{14} \text{ kg m}^{-3}$.
- **Inner Crust**, 9.7 to 10.3 km:
Consists of somewhat heavier nuclei such as ^{118}Kr , a superfluid of free neutrons, and relativistic degenerate electrons. The bottom is the site where nuclei dissolve at $\rho = \rho_{\text{nuc}} \approx 2 \times 10^{17} \text{ kg m}^{-3}$.
- **Interior (or Outer Core)**, 0 to 9.7 km:
A fluid mixture of superfluid free neutrons, a smaller number of superfluid superconducting free protons, and a small % of relativistic free electron gas.
- **Core? (or Inner Core)**
There may or may not be a solid inner core. The central density of a NS of this mass and size is estimated to be $10^{18} \text{ kg m}^{-3}$.

“Chandrasekhar Limit” for Neutron Stars

The relationship of radius to mass in a roughly inverse one, similar to white dwarfs, so they shrink and become more dense with increasing mass. The graphs³⁵¹ to the right illustrate the mass-radius relationship as computed by various “equations of state” (EOSs) for stars “containing nucleons and, in some cases, hyperons”. The EOSs will not be further described here. Clearly, however, radius shrinks as mass rises, but in a complex manner.

Just as a WD collapses to a NS when electron degeneracy can no longer support it, a NS has a maximal mass limit (the Chandrasekhar Limit for Neutron Stars) beyond which neutron degeneracy pressure can no longer support it. This mass limit is $2.2 M_\odot$ if the star is not rotating (unlikely) and $2.9 M_\odot$ if the star is rotating rapidly (as a result of centrifugal force effects). When the star exceeds these limits and becomes dynamically unstable, it will collapse to a black hole.

Although *IMA2* p. 583 uses the term “*The Chandrasekhar Limit for Neutron Stars*”, this limit is not universally called this name. Often it is unnamed, or named according to the models employed—including the original **Tolman**—



³⁵⁰ http://en.wikipedia.org/wiki/Neutron_star

³⁵¹ Lattimer JM and Prakash M, “Neutron Star Structure and the Equation of State”, *The Astrophysical Journal*, 550:426-442, March 20, 2001, <http://arxiv.org/abs/astro-ph/0002232>

Oppenheimer–Volkoff limit.³⁵² The equations of state for neutron stars are evolving and newer models will continue to appear.

Rotation and Angular Momentum

The slowly rotating stellar core becomes a rapidly rotating NS. The period of the NS is estimated as follows (based on conserved angular momentum and the change in moment of inertia):

$$P_{\text{NS}} \approx 3.8 \times 10^{-6} P_{\text{core}}$$

It is difficult to estimate the rate of rotation of a progenitor stellar core, but *IMA2* p. 584 uses a value of 1350 s as representative. This leads to a NS period of 5×10^{-3} s (5 msec).

Neutron stars slow in rotation at a typical slow down rate of 10^{-15} seconds per rotation.³⁵³

Magnetic Field

Neutron stars are expected to have intense magnetic fields because the field lines of a collapsing stellar core are frozen in. The field strength therefore increases markedly compared to the stellar cores from which NSs form. For spherically symmetrical objects, the magnetic field strength will be inversely proportional to the surface area. With marked contraction, the field strength from the frozen-in field lines is greatly increased. If an iron stellar core were to have a field strength of $B \approx 5 \times 10^4$ T (a maximum observed for free WDs, admittedly not the same starting point), the resulting NS field strength is 1.3×10^{10} T. The maximum B observed is 10^{11} T (10^{15} G) in what are called **magnetars**.³⁵⁴ Typical NS field strength $B = 10^8$ K or less. A formula for B is given below.

Neutron Star Temperatures and Emissions

Neutron stars begin very hot from their origin in a supernova. Their initial temperature of $\sim 10^{11}$ K rapidly declines by a process that emits profuse neutrino radiation, the Urca process (named in 1941 after a casino in Rio de Janeiro that rapidly depleted one's money, the Cassino da Urca). The neutrinos are formed in great numbers and carry away much energy. In one day, the T drops to about 10^9 K. Interior T is 10^8 K after a few hundred years, at which time the surface T has dropped to several million K. Surface T will remain at $\sim 10^6$ K for $\sim 10,000$ years. At this T_{eff} , the Stefan–Boltzmann law predicts a luminosity of 7.13×10^{25} W, about 1/5 of solar luminosity (L_{\odot} = of 3.846×10^{26} W). The Planck distribution peak by Wien Displacement occurs at a wavelength of 2.9 nm, corresponding to a photon energy of 0.4 keV, thus predominantly soft X-ray emissions. This range is best studied by space-based X-ray observatories such as the Chandra X-ray Observatory (1999–), ROSAT (Röntgensatellit, 1990–1998), ASCA (1993–2001) and XMM-Newton (1999–), etc.

Miscellaneous

Gravitational light deflection arises at a neutron star. Due to this GR light deflection, substantially more than half of the spherical surface is visible, an effect that increases with increasing mass.³⁵⁵

The weak visible light emitted is presumably perceived, if at all, as white or bluish-white (in view of the peak emissions in soft X-ray).

There are about 2000 known neutron stars in the MW and Magellanic Clouds (per Wikipedia).

³⁵² http://en.wikipedia.org/wiki/Tolman%20%93Oppenheimer%20%93Volkoff_limit

³⁵³ http://en.wikipedia.org/wiki/Neutron_star

³⁵⁴ <http://www.cv.nrao.edu/course/astr534/Pulsars.html>

³⁵⁵ <http://www.spacetimetavel.org/galerie/galerie.html>

Pulsars

The word *Pulsar* is a contraction of *pulsating star*. The pulsations of rotating neutron stars with strong magnetic fields were first detected on November 28, 1967 by Jocelyn Bell Burnell (1943–) and Antony Hewish (1924–). It was a signal that recurred every sidereal day (now known to be arising from pulsar PSR 1919+21, where *PSR*=*Pulsating Source of Radio*). The pulse period P was 1.337 s (0.75 revolutions s⁻¹ or Hz). Although PSR 1919+21 emits in radio wavelengths, pulsars have subsequently been found that emit in infrared, visible light, X-ray, and/or gamma ray wavelengths.³⁵⁶ Geminga (PSR J0633+1746) only radiates pulses in soft X-rays and gamma rays (with period 237 ms)—it is radio quiet. It has been identified in visible light, but I am unclear whether there are visible light pulsations. (IMA2 p. 590 and *here*³⁵⁷)

Most known pulsars are in the MW galaxy or in nearby galaxies. There were about 1500 known pulsars as of 2005 (IMA2 p. 586, updated to 2008 in May 2012 according to ATNF³⁵⁸ and 2123 according to SIMBAD.³⁵⁹ They have periods mostly between 0.25 s and 2 s, with average pulsation period of 0.795 s (1.26 Hz). The longest known pulsar period is 9.437 s (0.106 Hz)³⁶⁰ or PSR 1841-0456 with period=11.8 s and rotation at 0.085 Hz (IMA2 p. 588 and *here*,³⁶¹ showing this star is subject to glitches).

Pulsar PSR J1748-2446ad (Terzan 5ad), discovered in 2005, is, as of 2010, the swiftest spinning pulsar currently known, spinning at 716 times a second (716 Hz), therefore with a period=1.4 msec.³⁶²

An X-ray pulsar, IGR J00291+593, spins at 599 s⁻¹ (599 Hz), thus with a period of 1.7 msec.

Pulsars, like other neutron stars, usually have a highly regular rate of rotation. The period normally increases at a rate of about

$$\dot{P} \equiv dP/dt \approx 10^{-15},$$

(a unitless result from s⁻¹/s⁻¹). If \dot{P} were to remain constant, the characteristic lifetime would be the time required to bring the rotation to a halt, and calculated values range from a few tens of millions of years to about 500 million years.

Pulsars as Rapidly Rotating Neutron Stars

Thomas Gold (1920–2004) first concluded that these pulsations must be due to rapid rotation of a neutron star.

Possible mechanisms considered for pulsating stars with this short period include Binary Stars (inconsistent with WDs and with binary NSs due to General Relativity), Pulsating or Oscillating WDs or NSs (infeasible mathematically), and Rotating Stars (the favored explanation).

There is substantial evidence supporting an origin of NS from core-collapse supernovae. The NS often have a high peculiar velocity, and few are found in binary systems. This may indicate that an asymmetrical explosion or a jet propelled the NS from its starting point. By running the clock backwards, it appears the Crab Nebula has accelerated in its expansion since its inception.

If a NS rotates too fast, it will be torn apart. The **minimum rotation period** a NS must have for which gravity still overcomes centrifugal effects is given by (IMA2 p. 590):

³⁵⁶ <http://en.wikipedia.org/wiki/Pulsar>

³⁵⁷ <http://en.wikipedia.org/wiki/Geminga>

³⁵⁸ <http://www.atnf.csiro.au/research/pulsar/psrcat/> click on “Table” without filling in any blanks

³⁵⁹ Use query criteria: otype = ‘psr’

³⁶⁰ Young MD et al, “A radio pulsar with an 8.5-second period that challenges emission models” *Nature* 400, 848-849 (26 August 1999) doi:10.1038/23650. This describes PSR J2144-3933

³⁶¹ <http://iopscience.iop.org/0004-637X/673/2/1044/fulltext/>

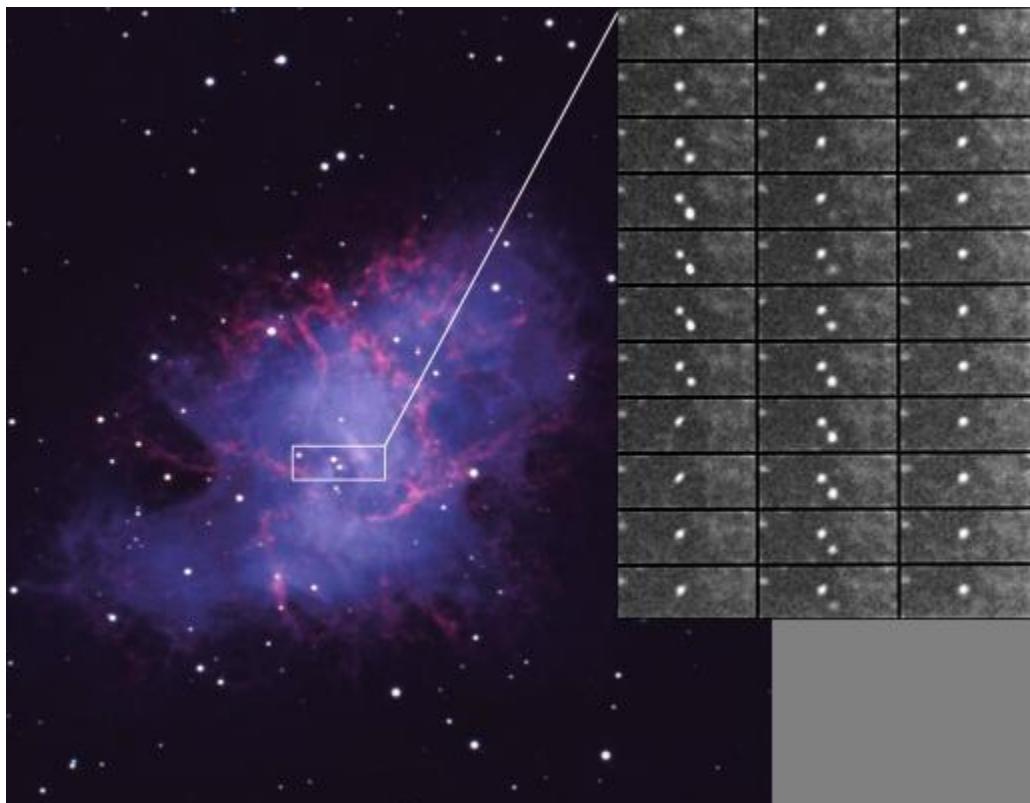
³⁶² Hessels JWT, et al, “A Radio Pulsar Spinning at 716 Hz”, *Science* 31 March 2006, Vol. 311 no. 5769 pp. 1901-1904, DOI: 10.1126/science.1123430

$$P_{\min} = 2\pi \sqrt{\frac{R^3}{GM}}$$

For a $1.4 M_{\odot}$ NS, the minimum allowed period is 5×10^{-4} s = 0.5 msec. All pulsars to date have had periods exceeding this lower limiting value and are therefore not in immediate risk of flying apart.

Many pulsars have been associated with supernova remnants, including the Crab Pulsar (PSR B0531+21 aka PSR J0534+2200, rotation period ~33 msec) and the Vela Pulsar (PSR J0835-4510 aka PSR B0833-45, radio period ~89 msec). The Crab Pulsar's rotation rate is incompatible with a WD (as is Vela's). (*IMA2* p. 590)

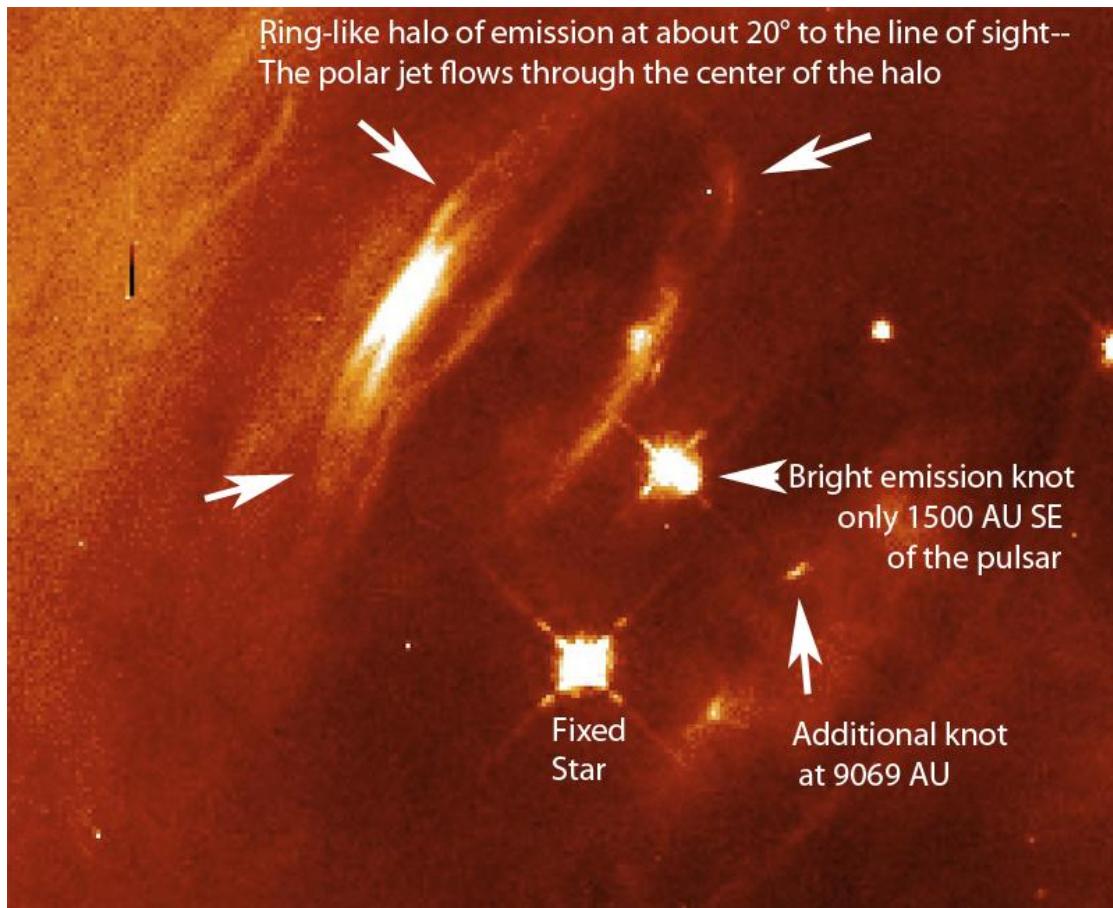
Despite the precise measurements that have been made of rotation periods, pulsars do not always maintain a steady and predictable increase in period. (For this reason, I believe that they will never make good astronomical time-keeping devices.) Young pulsars, including the Vela and Crab pulsars, are subject to "glitches" (see below).



Pulsation of Crab Pulsar: "A time sequence for the pulsar in the Crab nebula is shown, and its location in the Crab nebula (as shown in a KPNO 4-meter Mayall image [1973]). The images of this sequence were also obtained with the 4-meter Mayall telescope on Kitt Peak, during the night of 20 October 1989 using a standard B-band optical filter and the Kitt Peak Photon Counting Array (KPCA). Phased accumulation over almost 2 hours was necessary to create this image; the observed period that day was 33.36702 milliseconds. Each of the 33 images represents a time slice of about 1 millisecond in the pulsar period. The brighter, primary pulse is visible in the first column: the weaker, broader inter-pulse can be seen in the second column." (See graph next page of pulse sequences.) Sequence of visible light images are arranged from top to bottom, then left to right. Foreground star is invariant while Crab Pulsar varies in light output. (See light curve.) (N. A. Sharp / AURA / NOAO / NSF)³⁶³

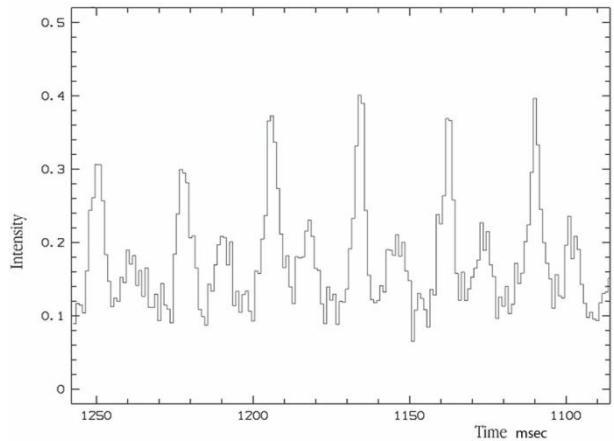
³⁶³ http://messier.seds.org/more/m001_pulsar.html

The following image also depicts the Crab Pulsar in the center of the Crab Nebula, but it is much higher in resolution (from HST). It shows a knot of hot emission very near the pulsar (slightly southeast of the pulsar, separated from it by only 1200 AU). It also shows “that in the direction opposite the knot, the Crab pulsar is capped by a ring-like ‘halo’ of emission tipped at about 20 degrees to our line of sight. In this geometry the polar jet flows right through the center of the halo. The newly discovered ring may mark the boundary between the polar wind and jet, and an equatorial wind that powers a larger torus of emission surrounding the pulsar.” (HST, Hester and Scowen, 1995):³⁶⁴



Crab Pulsar Light Curve: The period of the Crab Pulsar pulsations is measured between the higher peaks. “One full revolution of the neutron star corresponds to the distance from one high peak to the next, and the diagram therefore covers six consecutive revolutions (about 200 milliseconds).”³⁶⁵ The weaker, broader inter-pulse peaks are also evident (image to right)

The Crab Pulsar period has varied from 33.35 ms in May 1988 to 33.67 ms in May 2012,³⁶⁶ with a few glitches along the way. (Period in msec is given by $1000/(N \mu \text{Hz})$.)



³⁶⁴ Hester et.al., “WFPC2 Studies of the Crab Nebula. I. HST and ROSAT Imaging of the Synchrotron Nebula.” *Astrophysical Journal*, Vol. 448, p. 240-263 and plate 10-15 (July 20, 1995) Image obtained from http://messier.seds.org/more/m001_hst.html#hester1995 Also found at *IMA2* p. 594.

³⁶⁵ <http://www.eso.org/public/images/eso9948i/> including image of light curve

³⁶⁶ <http://www.jb.man.ac.uk/pulsar/crab/crab2.txt>

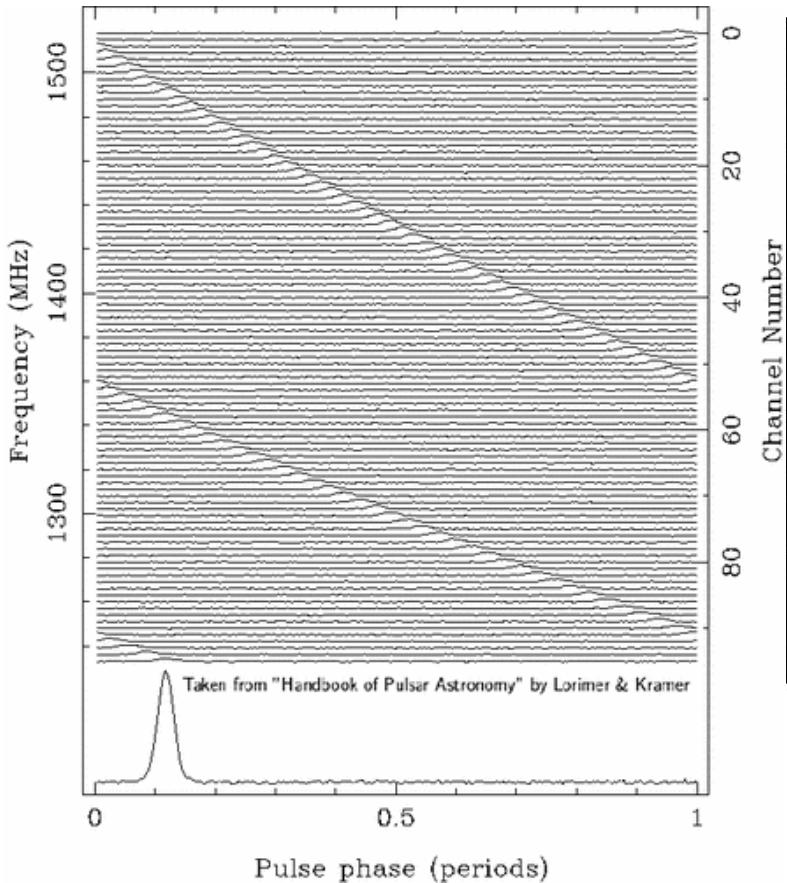
Synchrotron Radiation and Curvature Radiation

The Crab Nebula is connected with its central pulsar. The ghostly glow the nebula exhibits was proposed (by Iosif Shklovsky, 1916-1985) to be synchrotron radiation. Synchrotron radiation is generated by acceleration of relativistic electrons as a result of the Lorentz force \mathbf{F} (where $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$) exerted in the presence of a magnetic field \mathbf{B} . The resulting acceleration causes them to circle or spiral around often curving magnetic field lines. \mathbf{E} is negligible here. The acceleration of charged particles in a magnetic field leads to emission of synch. rad according to the right hand rule (where v=index finger, B=middle finger, F=thumb). (A related radiation, **curvature radiation**, arises from electrons that are moving along the curving field lines.) The “non-thermal” emission is a strongly linearly polarized but otherwise featureless broad spectrum continuum with energy distribution which depends on the velocity distribution of the electrons, and which is quite distinct from blackbody (aka “thermal”) radiation. For relativistic electron velocities, the emission pattern is sharply collimated forward (relativistic beaming) in the direction of the electron’s current velocity vector. Parts of the Crab Nebula emit light which is 60% linearly polarized, confirming its origin as synchrotron radiation.

The relativistic electrons and the synchrotron radiation should have long ago died away in the Crab Nebula if they were simply residuals from the SN. Instead, it appears that the central pulsar is replenishing both the relativistic electrons and the magnetic field. The calculated amount of energy being injected is a remarkable 5×10^{31} W or $10^5 L_\odot$. The NS is acting as a huge flywheel with a high amount of rotational kinetic energy, energy gradually imparted to the surrounding nebula. Using previously stated average values for a $1.4 M_\odot$ NS, the expected kinetic rotational energy loss resulting from the documented slowing of the pulsar matches well the energy needed to replenish the nebula. The required energy is not supplied by the radio pulses in the Crab—they are much too weak, only about 10^{24} W. The wisps of the Crab Nebula are moving outward at velocities as high as 0.5c. (IMA2 p. 595)

Pulse Characteristics

The radio pulses are usually received at 20 MHz to 10 GHz. Passage of radio waves through the ISM (or nebular media) causes the ISM electrons to oscillate, the net effect being that the waves are retarded somewhat below vacuum c, with greater retardation at lower frequencies. This causes a spreading out or *dispersion* of the pulses. (See diagram to right, which shows “Uncorrected dispersive delays for a pulsar observation over a bandwidth of 288 MHz (96 channels of 3 MHz width each), centered at 1380 MHz. The delays wrap since the data are folded (i.e. averaged) modulo the pulse period.”).³⁶⁷ Note that this depicts dispersion by frequency, not waveform drift with time or subpulses. The amount of dispersion may be used to estimate distance to the pulsar (at best to ~30% $\Delta d/d$ accuracy). “Since pulsar observations almost always cover a wide bandwidth, uncorrected differential delays across the band will cause dispersive smearing of the pulse profile. For pulsar searches, the DM is unknown and becomes a search parameter much like the pulsar spin frequency. This extra search dimension is one



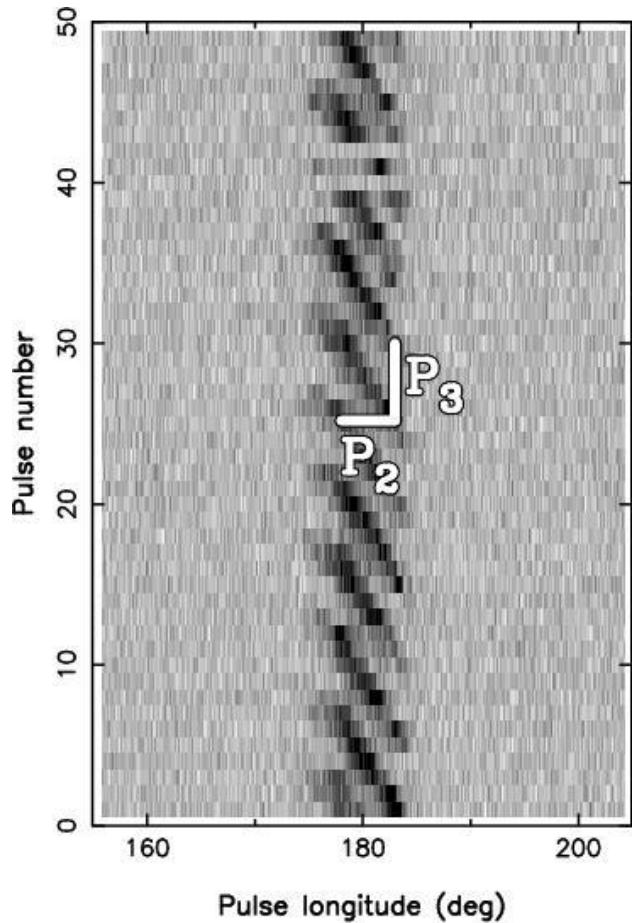
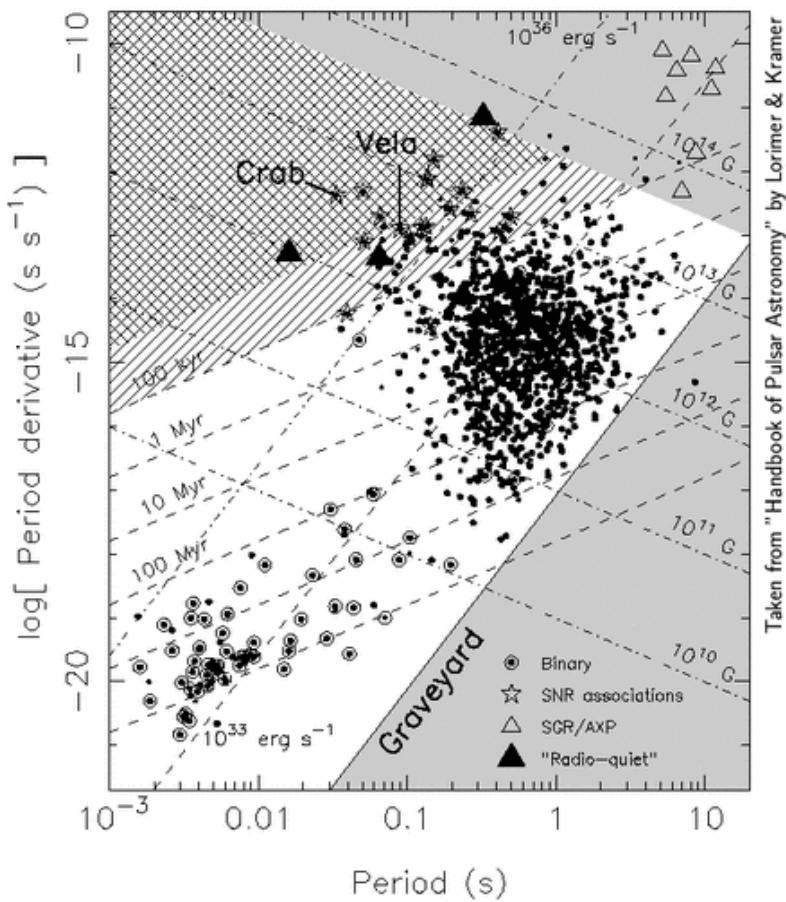
³⁶⁷ Pulse Dispersion:

- all quotes from <http://www.cv.nrao.edu/course/astr534/Pulsars.html>,
- diagram from <http://www.jb.man.ac.uk/research/pulsar/handbook/figures/med/1.08.gif>
- see also <http://astronomy.swin.edu.au/cms/astro/cosmos/p/Pulsar+Dispersion+Measure>

of the primary reasons that pulsar searches are computationally intensive.”

Despite the Crab Nebula light curve, pulses in general are very brief and occupy only a small part of the period, typically 1 to 5% in the 20 MHz to 10 GHz band. The individual pulses vary somewhat, but the averaged waveform from integrating 100 pulses tends to be stable. Pulsars may have more than one wave pulse mode and switch between these pulse modes. There may be subpulses, including ones that “drift”, and the different modes may have different subpulses, up to 5 or more. (IMA2 p. 598). Pulses may disappear (“null”) for 100 periods or so, then return. They are strongly linearly polarized.

The graph to the right shows “a pulse-stack of fifty successive pulses of PSR B0818-13. Two successive drift bands are vertically separated by P3 [measured in periods] and horizontally by P2 [measured in phase change].”³⁶⁸



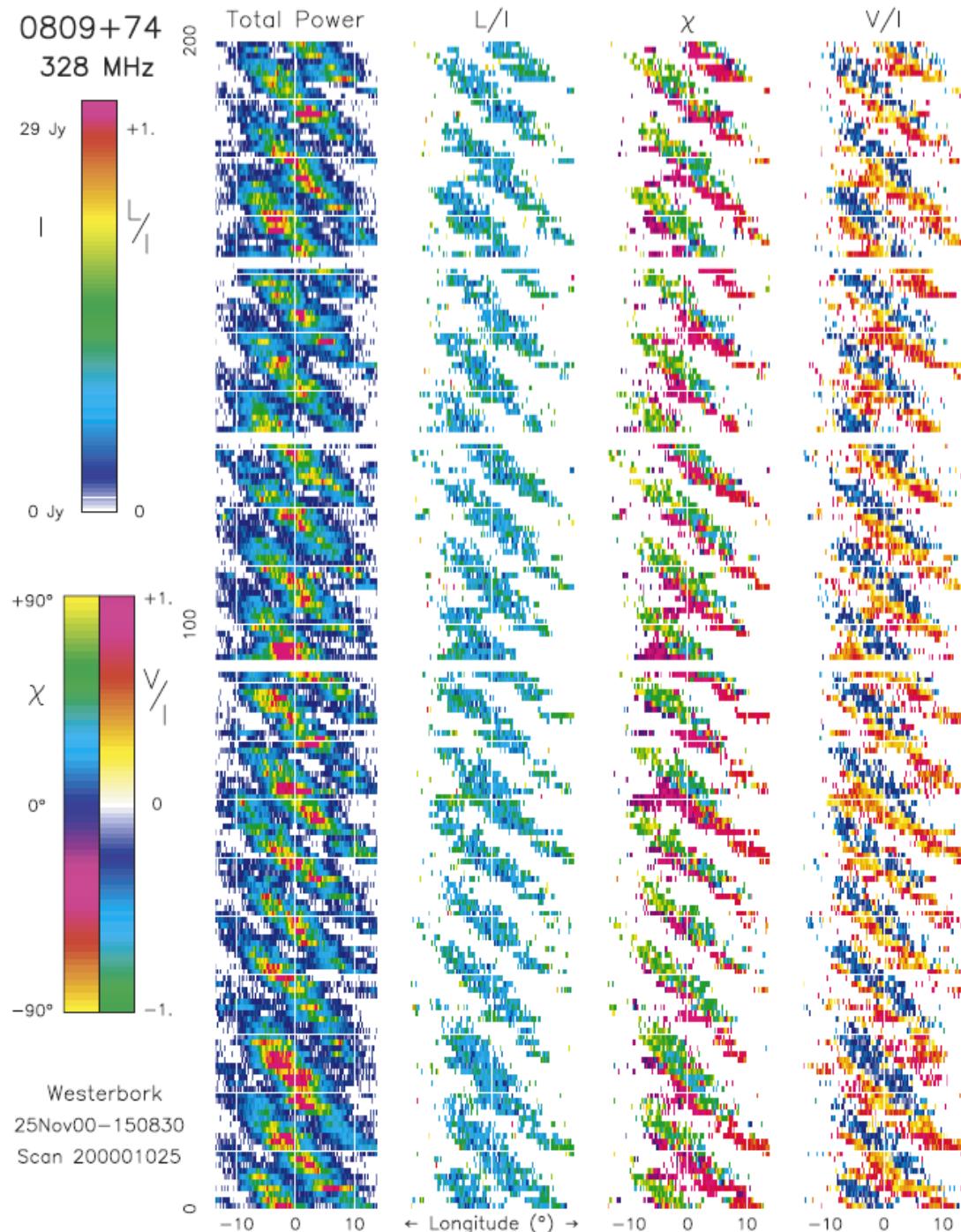
Pulsars may be grouped with a graph of period versus period time derivative, the so-called $P - \dot{P}$ diagram (see graph on left).³⁶⁹ This “is useful for following the lives of pulsars, playing a role similar to the Hertzsprung-Russell diagram for ordinary stars. It encodes a tremendous amount of information about the pulsar population and its properties, as determined and estimated from two of the primary observables, P and \dot{P} . Using those parameters, one can estimate the pulsar age, magnetic field strength B , and spin-down power \dot{E} .³⁷⁰ Here, SNR=supernova remnant, SGR=Soft Gamma Repeaters, AXP=Anomalous X-ray pulsars.

³⁶⁸ Weltevrede, P., “The subpulse modulation properties of pulsars at 92 cm and the frequency dependence of subpulse modulation” *Astronomy and Astrophysics*, Volume 469, Issue 2, July II 2007, pp.607-63, DOI: 10.1051/0004-6361:20066855

³⁶⁹ <http://www.jb.man.ac.uk/research/pulsar/handbook/figures/med/1.13.gif>

³⁷⁰ www.cv.nrao.edu/course/astr534/Pulsars.html (Good review of pulsar properties)

The following rather complex colored graph depicts stacked pulse graphs for Pulsar PSR B0809+74 at 328 MHz, and explores the distribution of polarization:³⁷¹



"Color display of a 200-pulse portion... The first column of the displays [on the left] gives the total intensity (Stokes parameter I [total power]), with the vertical axis representing the pulse number and the horizontal axis pulse longitude [relative phase], colour-coded according to the ... top bar to the left of the displays. The second and third columns give the corresponding fractional linear polarisation L/I ... and its angle χ ... The last column gives the fractional circular polarisation (V/I)...". Note the variation in pulse peak total power, and the drift in pulse phase. "The 'drifting' subpulses ... exhibit strikingly linear and circular polarization which appears to reflect the characteristics of two nearly orthogonally polarized emission 'modes'."

³⁷¹ Ramachandran R et al, "Pulsar 'drifting'-subpulse polarization: No evidence for systematic polarization-angle rotations", *A&A* 381, 993-999 (2002), DOI: 10.1051/0004-6361:20011525

Pulsar Model and Postulated Emissive Mechanism

The traditional model is of a rotating neutron star in which the axis of rotation differs significantly from the magnetic dipole field axis:

Diagram³⁷² to right depicts the traditional magnetic dipole model of a pulsar and its generation of the radio beam (emission cone). See discussion below.

The emissive mechanism is poorly understood, but a model is described as follows (from *here*³⁷³ and with my interpolations in [])

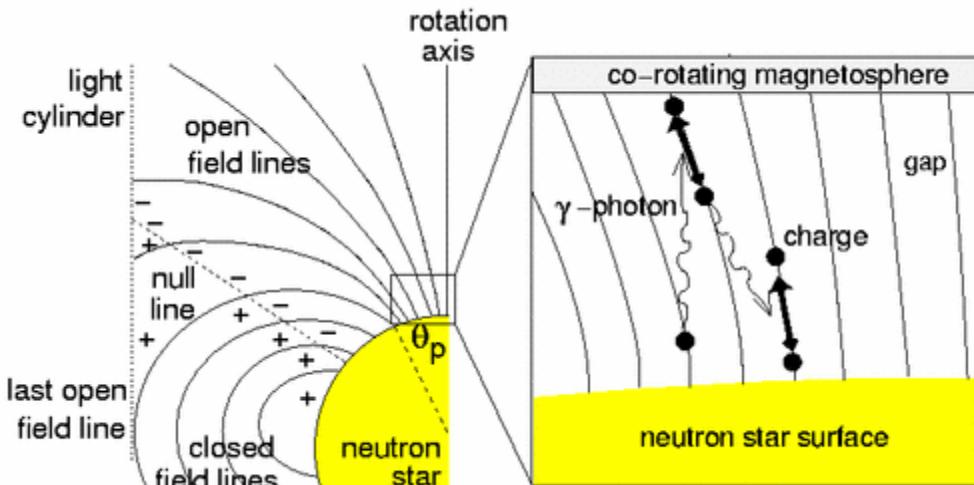
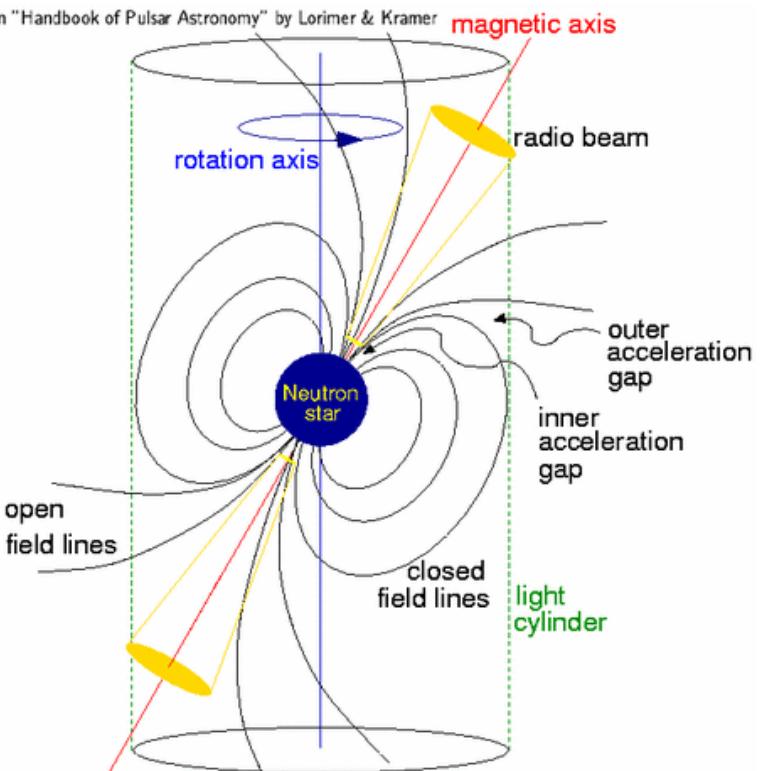
"The radio pulses originate in the pulsar magnetosphere. Because the neutron star is a spinning magnetic dipole, it acts as a unipolar generator. The total Lorentz force acting on a charged particle is [for SI units, $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$]."

"Charges in the magnetic equatorial region redistribute themselves by moving along closed field lines until they build up an

electrostatic field large enough to cancel the magnetic force and give $|\mathbf{F}|=0$. The voltage induced is about 10^{16} V in MKS units. However, the co-rotating field lines emerging from the polar caps cross the **light cylinder** (the cylinder centered on the pulsar and aligned with the rotation axis at whose radius the co-rotating speed equals the speed of light) and these field lines cannot close. Electrons in the polar cap are magnetically accelerated to very high energies along the open but curved field lines, where the acceleration resulting from the curvature causes them to emit *curvature radiation* that is strongly polarized in the plane of curvature. As the **radio beam** [of curvature radiation] sweeps across the line-of-sight, the plane of polarization is observed to rotate by up to 180 degrees, a purely geometrical effect. High-energy photons produced by curvature radiation interact with the magnetic field and lower-energy photons to produce **electron-positron**

pairs that radiate more high-energy photons. The final results of this cascade process are bunches of charged particles that emit at radio wavelengths. (Some of these steps are depicted in the diagram to the left.)³⁷⁴

Taken from "Handbook of Pulsar Astronomy" by Lorimer & Kramer



Taken from "Handbook of Pulsar Astronomy" by Lorimer & Kramer

The *death line* in the $P - \dot{P}$ diagram [beyond which lies the "graveyard"] corresponds to neutron stars with sufficiently low B and high P that the curvature radiation near the polar surface is no longer capable of generating particle

³⁷² <http://www.jb.man.ac.uk/research/pulsar/handbook/ch3.html>

³⁷³ www.cv.nrao.edu/course/astr534/Pulsars.html All quotes in this section are from this NRAO webpage except as noted.

³⁷⁴ <http://www.jb.man.ac.uk/research/pulsar/handbook/figures/med/3.03.gif>

cascades.” [In other words, at this point radio pulsations cease.]

“The extremely high brightness temperatures are explained by coherent radiation. The electrons do not radiate as independent charges e ; instead bunches of N electrons in volumes whose dimensions are less than a wavelength emit in phase as charges Ne . Since Larmor's formula [for radiation from a rotating electric dipole] indicates that the power radiated by a charge q is proportional to q^2 , the radiation intensity can be N times brighter than incoherent radiation from the same total number N of electrons. Because the coherent volume is smaller at shorter wavelengths, most pulsars have extremely steep radio spectra. Typical (negative) pulsar spectral indices are $\alpha \approx 1.7$ ($S \propto v^{-1.7}$), although some can be much steeper ($\alpha > 3$) and a handful are almost flat ($\alpha < 0.5$).”

The energy ($J\ s^{-1}$) emitted due to a rotating magnetic dipole is given by (IMA2 p. 597):

$$\frac{dE}{dt} = -\frac{32\pi^5 B^2 R^6 \sin^2 \theta}{3\mu_0 c^3 P^4}$$

The magnetic field at the pole of a NS is given by (IMA2 p. 599):

$$B = \frac{1}{2\pi R^3 \sin \theta} \sqrt{\frac{3\mu_0 c^3 I \dot{P}}{2\pi}}$$

where I is the moment of inertia, R = star radius, θ = angle between rotation axis and magnetic axis.

The radius of the light cylinder is given by $R_c = c/\omega = cP/2\pi$.

The strong electric field continuously rips electrons and positive ions from the surface, easily overcoming gravity, causing the rotating magnetosphere to be filled with charged particles. At the light cylinder, the particles cannot keep up with the co-rotation (which is at the speed of light) and are spun away as part of a pulsar wind that carries away part of the magnetic field as well. This kind of wind may explain the mechanism for replenishment of the Crab Nebula's energy. (IMA2 p. 601)

According to the model described in IMA2 p. 601, which differs only a little from the NRAO description, electrons and ions in the polar region are accelerated to relativistic speeds by the induced *electric field*. As the electrons follow the curved magnetic field lines, they emit curvature radiation that is relativistically formed into a narrow beam. Gamma ray photons produced by curvature radiation can spontaneously produce electron-positron pairs that subsequently are accelerated again radiate more high-energy photons, and the cascading process continues. Coherent beams of curvature radiation emitted by bunches of particles may be responsible for the individual subpulses. The narrowly directed radio wave beams emitted near the polar region sweep around the star light a lighthouse beacon might. A terrestrial detector detects the pulses as they rotate past the Earth.

Glitches are very sudden though infrequent changes in pulsar period. One seen for Vela in 2010 exhibited a relative change in frequency $\Delta v/v$ of 1.94×10^{-6} .³⁷⁵ Glitches may be a result of “starquakes” resulting from fractures, faulting, and shifting in the rigid solid crust, or perhaps an unpinning of superfluid vortices in the core of the NS, but this is a subject requiring more research. (IMA2 p. 602)

Fate of a Pulsar

It is unclear whether the magnetic field must decay with time. In one view, this decay would take place with a characteristic time of perhaps 9 million years, which will lead eventually to cessation of detectable pulses. Alternatively, the magnetic field may not decay. But in either case, the rotation rate will slowly decrease and this will cause the beam to become weaker even if the magnetic field does not change. In the $P - \dot{P}$ diagram, they will enter the “graveyard” and become a radio-quiet NS. (IMA2 p. 602)

³⁷⁵ <http://www.hartrao.ac.za/news/100731vela/index.html> See also IMA2 p. 591.

Soft Gamma Repeaters

It is possible that magnetars are responsible for soft gamma repeaters SGRs. These objects emit hard X-rays and soft gamma rays up to 100 keV. Only a few are known, and these mostly correlate with SNRs. One detected in August 2008 by a giant X-ray outburst, is SGR 0501+4516. It “is estimated to lie about 15 000 light-years away, and was undiscovered until its outburst gave it away”³⁷⁶ The sudden energy outburst of an SGR is thought to be caused by a crack in the crust that allows a release of energy greatly exceeding the Eddington luminosity limit for stars, though it is very short-lived. (*IMA2* p. 603) SGR 0501+4516 was observed with XMM-Newton and Swift-XRT quasicontinuous monitoring and its pulses were fitted to blackbody radiation and power law curves. The data strongly argue that this source is a magnetar candidate with a magnetic field of $B \sim 2 \times 10^{10}$ T. “The discovery of SGR0501+4516 , and its AXP-like characteristics, represents another piece of evidence in the unification of the magnetar candidate class, weakening further the differences between AXPs, TAXPs, and SGRs.”³⁷⁷ Here, AXPs = Anomalous X-ray Pulsars, and TAXPs = Transient Anomalous X-ray Pulsars. SGRs may arise in young SNRs (*IMA2* p. 603) and may be relatively short-lived. However, at least some AXPs are not associated with SNRs. At least one SGR has been described for which the surface magnetic field is much lower than the usual range for magnetars, with B not greater than 7.5×10^{12} gauss (7.5×10^8 T)³⁷⁸

³⁷⁶ Rea N, et al, “The first outburst of the new magnetar candidate SGR 0501+4516”, <http://arxiv.org/abs/0904.2413v1>

³⁷⁷ ibid.

³⁷⁸ Rea N., et al, “A Low-Magnetic-Field Soft Gamma Repeater”, *Science* 12 November 2010, Vol. 330 no. 6006 pp. 944-946, DOI: 10.1126/science.1196088

General Relativity and Black Holes (Chapter 17, mostly omitted)

"*Spacetime tells matter how to move; matter tells spacetime how to curve.*"³⁷⁹

This chapter has not yet been read in my astronomy courses and is not summarized here, aside from an excerpt below.

Schwarzschild radius

The Schwarzschild radius is named for Karl Schwarzschild (1873–1916),³⁸⁰ who derived this first exact solution in the Einstein field equations of general relativity. It defines the **event horizon** of a black hole. For a non-rotating star or black hole, it is given by

$$R_S = \frac{2GM}{c^2}$$

Black holes are fully described by the mass, angular momentum, and electric charge—externally at least, a black hole "has no hair" (*IMA2* p. 634–640). The maximum angular momentum for a black hole is

$$L_{\max} = \frac{GM^2}{c}$$

"Rotating black holes are surrounded by a region of spacetime in which it is impossible to stand still, called the **ergosphere**. This is the result of a process known as frame-dragging; general relativity predicts that any rotating mass will tend to slightly "drag" along the spacetime immediately surrounding it. Any object near the rotating mass will tend to start moving in the direction of rotation. For a rotating black hole this effect becomes so strong near the event horizon that an object would have to move faster than the speed of light in the opposite direction to just stand still... The ergosphere of a black hole is bounded by the (outer) event horizon on the inside and an oblate spheroid, which coincides with the event horizon at the poles and is noticeably wider around the equator. The outer boundary is sometimes called the **ergosurface**. Objects and radiation can escape normally from the ergosphere. Through the **Penrose process**, objects can emerge from the ergosphere with more energy than they entered. This energy is taken from the rotational energy of the black hole causing it to slow down."³⁸¹

For a rotating black hole, the event horizon (like the ergosphere) is oblate and wider at the equator. For a maximally rotating black hole, the event horizon at the equator is (*IMA2* p. 641):

$$r = \frac{R_S}{2} = \frac{GM}{c^2}$$

³⁷⁹ Wheeler JA, Ford KW, *Geons, Black Holes, and Quantum Foam: A Life in Physics*, p. 235, 1998.

³⁸⁰ http://en.wikipedia.org/wiki/Karl_Schwarzschild

³⁸¹ http://en.wikipedia.org/wiki/Black_hole

Close Binary Star Systems (Chapter 18)

Gravity Potentials in a Close Binary Star System

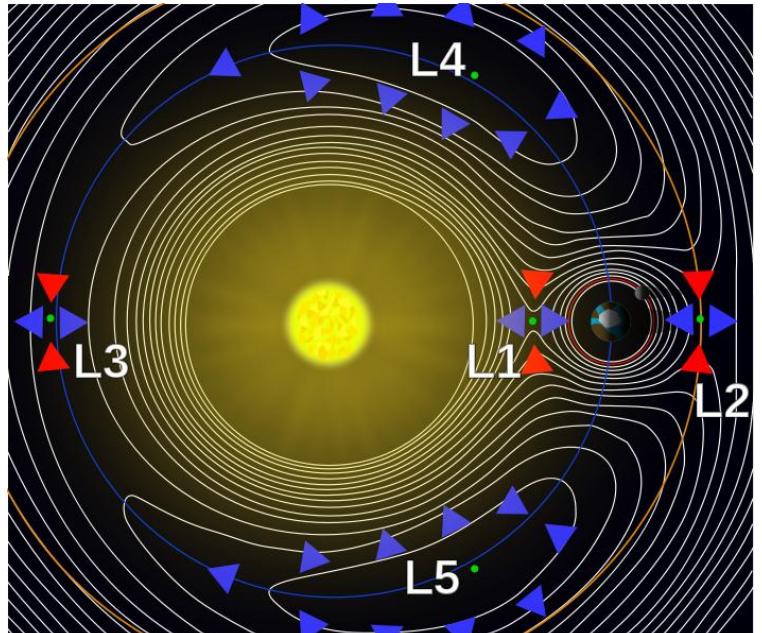
Binary systems appear to be among the astrophysicist's best friends. Binary systems are common, but in most, the two components are widely spaced and negligibly interact. In a **close binary system**, the stars are sufficiently close that at least one gravitationally distorts the other's atmosphere—their separation is roughly equal to the diameter of the larger star. A rotating star that is tidally deformed pulsates and dissipates energy as heat. Such a system may eventually reach a state of minimum energy for its angular momentum, resulting in **circulated orbits** and, at least for normal sized non-compact stars, **synchronous rotation** (in which, like the Earth's moon, the same hemisphere always faces the same hemisphere of the other star). In this situation, no further energy is lost to tidal interactions. (IMA2 p. 653 and *here*³⁸²)

Reference Frame: When referring to orbital rotation in a binary system, in all cases each body actually orbits around the barycenter or center of gravity of the system. It is convenient to analyze binary system motion using a co-rotating coordinate system with the same period as the two co-orbiting bodies and centered on the barycenter (center of mass).

Equipotential Surfaces and Lagrangian Points

In the diagram to the right in a rotating frame of reference, a Sun-like star is depicted at the center orbited by an Earth-like planet having an orbiting satellite moon. The actual barycenter is not shown. The red arrowheads indicate "downhill" gradients (i.e., the direction in which gravitational potential energy U for a test mass decreases to a more negative value) and the blue arrowheads indicated uphill gradients. The force on a test mass is given by $\vec{F} = -m\nabla\Phi$, and it is always perpendicular to the equipotential contour.

The forces on a test object are strongest when the contours of the effective potential are closest together and weakest when the equipotential contours are far apart.



The diagram depicts the **5 Lagrangian points (Lagrange points)**: The five positions in an orbital configuration where a small object of negligible mass and affected only by gravity can theoretically be **stationary** relative to two larger objects (such as a satellite with respect to the Earth and Moon). The Lagrangian points mark positions where the combined gravitational pull of the two large masses provides precisely the centripetal force required to cause the small object to orbit with the larger masses about the barycenter.³⁸³ L₁, L₂ and L₃ are unstable, meaning that any small perturbations can dislodge an object at any of these.

The effective gravitational potential energy per unit mass (of a test object) is given by (IMA2 p. 655):

$$\Phi = -G \left(\frac{M_1}{r_1} + \frac{M_2}{r_2} \right) - \omega^2 r^2$$

At the precise L₁, L₂, and L₃ points, there is no net force on a test mass because $\frac{d\Phi}{dx} = 0$ (IMA2 p. 655). However, a test mass at these points is unstable (some like ESA say *metastable*, others *semi-stable*) because

³⁸² http://en.wikipedia.org/wiki/Binary_star

³⁸³ Lagrange (Lagrangian) Points:

- http://map.gsfc.nasa.gov/mission/observatory_12.html (from WMAP, which orbits about L₂, as will JWST)
- http://en.wikipedia.org/wiki/Lagrangian_point (color diagram shown)
- <http://map.gsfc.nasa.gov/media/ContentMedia/lagrange.pdf> (mathematical derivation)

they represent local maxima of Φ . More precisely, these points perch at the apex of a saddle, and almost all perturbations cause the mass to accelerate “downhill” more or less in the direction of one of the blue arrowheads.

The L₁ point is most relevant in the study of close binary systems. L₁ is also good for observing the Sun—as with SOHO, the Advanced Composition Explorer ACE, and the Global Geospace Science (GGS) WIND satellite. Approximate formulas for the distance from L₁ to M₁ and L₁ to M₂ are given at *IMA2* p. 656.

The L₂ point is a good location for placing a satellite to observe the universe while protected by the Earth from solar radiation (such as WMPA, JWST, Gaia, Herschel, and Planck).

The L₃ point would be concealed by the Sun and is not suitable for satellites.

The L₄ and L₅ points (triangular or Trojan Lagrange points) are located about 60 degrees ahead of or behind a secondary mass (like the Earth) orbiting a larger mass (like the Sun). “The centres of mass of M and m and the Lagrange point L₅ form an equilateral triangle with the edge length r.”³⁸⁴ where r is the length between the centers of mass of M the primary mass and m the secondary mass. They are at the apex of a gravitational potential hill falling off in all directions, and again with $\frac{d\Phi}{dx} = 0$. However, when perturbed, instead of accelerating away as might be expected, a test mass moving away from an L₄ or L₅ point is acted on by the Coriolis force, causing it to go into a stable circular orbit around the point. Therefore, an object at an L₄ or L₅ point is truly stable. They may someday be employed for space station locations. “[Natural] Objects found orbiting at the L₄ and L₅ points are often called **Trojans** after the three large asteroids Agamemnon, Achilles and Hector that orbit in the L₄ and L₅ points of the Jupiter-Sun system... There are hundreds of Trojan Asteroids in the solar system. Most orbit with Jupiter, but others orbit with Mars. In addition, several of Saturn's moons have Trojan companions. No large asteroids have been found at the Trojan points of the Earth-Moon or Earth-Sun systems. However, in 1956 the Polish astronomer Kordylewski discovered large concentrations of dust at the Trojan points of the Earth-Moon system. Recently, the DIRBE instrument on the COBE satellite confirmed earlier IRAS observations of a dust ring following the Earth's orbit around the Sun. The existence of this ring is closely related to the Trojan points, but the story is complicated by the effects of radiation pressure on the dust grains.”³⁸⁵ See NASA animation of the first known Earth Trojan, 2010 TK7.³⁸⁶

Roche Lobe, Roche Limit, and Roche Sphere

The **Roche lobe** (named for Édouard Albert Roche, 1820 – 1883) is the approximately cylindrically symmetrical 3D region of space around a star in a close binary system within which orbiting material is gravitationally bound to that star. The Roche lobe for close binaries has an approximately tear-drop 3D shape. If portions of the star expand past its Roche lobe (typically the atmosphere), then the atmospheric material can escape the gravitational pull of that star. It may pass into the “circumbinary” region that surrounds both stars, such as just inside the dumbbell shape that passes through L₂.³⁸⁷ But in a close binary system the atmospheric material can most easily escape one Roche lobe by passing in to the companion star’s Roche lobe through the inner Lagrangian point L₁. The apex of each Roche lobe’s tear-drop is at the L₁ Lagrangian point of the system.. (See diagram that follows, which nicely depicts the relatively low potential barrier that exists for passage of material through L₁ compared to other nearby “higher” routes.) The Roche lobe is different from the Roche limit (below) and the **Roche sphere**, which approximates the gravitational sphere of influence of one astronomical body in the face of perturbations from another heavier body around which it orbits.³⁸⁸

The **Roche limit** (less relevant to this discussion) is the distance or radius within which a secondary or satellite smaller celestial body M₂, held together solely by its own self-gravity, will disintegrate due to the distorting effect of the primary larger celestial body M₁. The latter exerts tidal forces exceeding M₂’s gravitational self-attraction. Tidal forces become destructive when portions of M₂ orbit closer to the M₁.

³⁸⁴ http://esamultimedia.esa.int/docs/edu/HerschelPlanck/EN_13e_L_Points_EarthMoonSystem.pdf

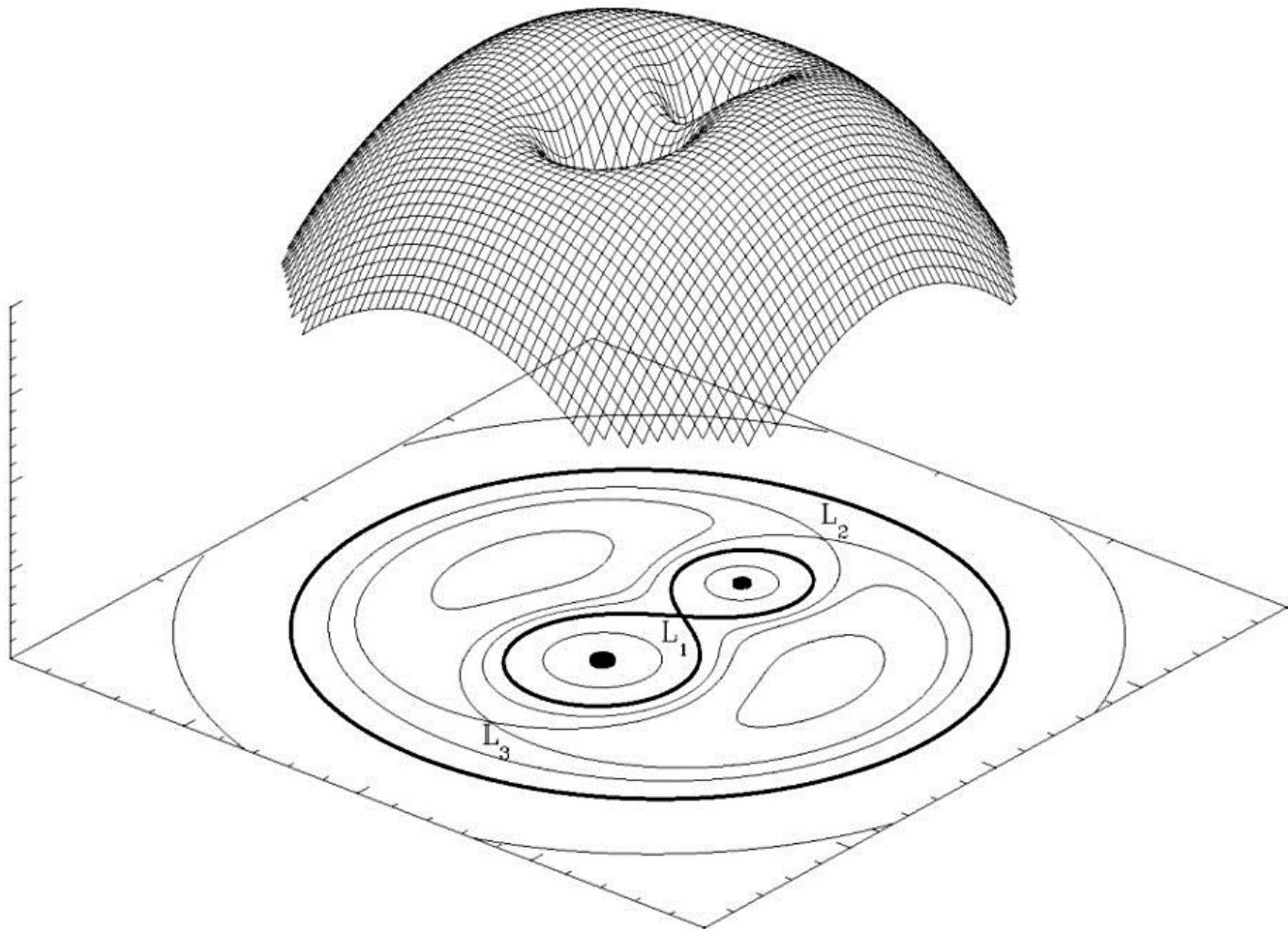
³⁸⁵ http://map.gsfc.nasa.gov/mission/observatory_l2.html

³⁸⁶ http://www.nasa.gov/multimedia/videogallery/index.html?media_id=103550791

³⁸⁷ <http://space.about.com/od/glossaries/g/rochelobe.htm>

³⁸⁸ http://en.wikipedia.org/wiki/Hill_sphere

primary attempt to move more quickly (orbit at higher angular velocity) than portions farther away. (IMA2 p. 723 and *here*³⁸⁹)



3D Equipotential surface showing potential wells for two close binary stars; also, a central plane cross section showing equipotential curves and lines, the tear-shaped Roche lobes separated by L_1 , the “dumbbell” passing through L_2 , and the L_3 Lagrangian point.³⁹⁰

Classes of Binary Star Systems

Close binaries are classified as **detached** (neither fills its Roche lobe), **semi-detached** (one fills its Roche lobe), or **contact** (both fill their Roche lobes).

Detached binary stars are widely separated (radii are much smaller than their separation), maintain a spherical shape, and have little influence on each other.

A **semi-detached** star in a binary system can expand beyond its Roche lobe and contribute atmospheric material to its companion. The overflowing star is usually called the **secondary** and labeled **M₂**. The companion star which has not filled its Roche lobe and which receives the material from M₂ is designated the **primary** or **M₁**.

Contact binaries share a common atmosphere bounded by a dumbbell-shaped equipotential surface and freely exchange atmospheric material.

³⁸⁹ http://en.wikipedia.org/wiki/Roche_limit

³⁹⁰ <http://hemel.waarnemen.com/Informatie/Sterren/hoofdstuk6.html#h6.2>

Mass Transfer Rate and Release of Gravitational Energy

The rate of mass transfer \dot{M} for a semi-detached binary with mass equaling its companion is roughly estimated at (IMA2 p. 658):

$$\dot{M} \approx \pi R d \rho \sqrt{\frac{3kT}{m_H}}$$

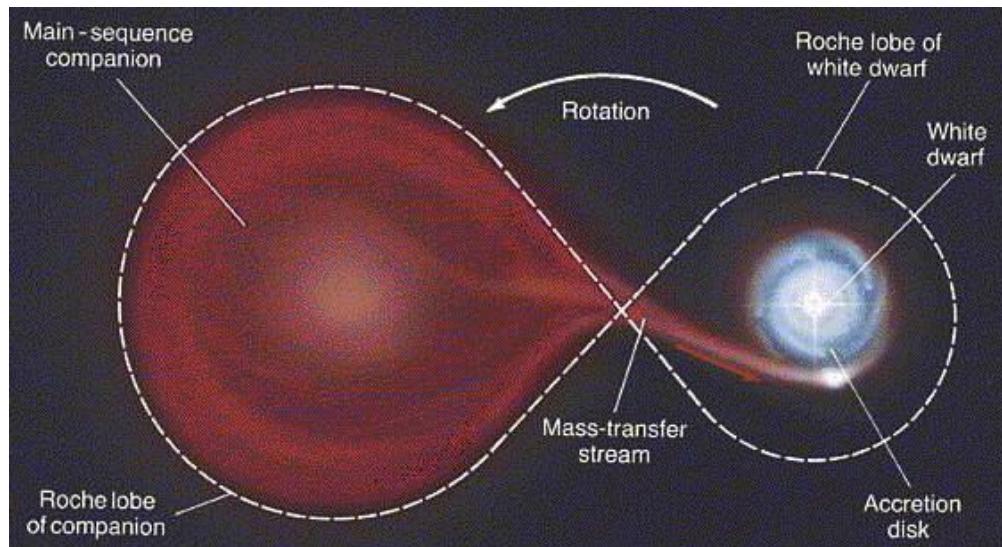
where R is the radius of the expanded star, d is the distance over which the two stars would overlap if they both had the radius of the expanded star, and m_H is the mass of a hydrogen atom. The more the stars overlap, the greater is d , ρ , and the T at the region of transfer, markedly increasing the transfer rate. Typical mass transfer rates for semidetached binaries range from 10^{-11} to $10^{-7} M_{\odot} y^{-1}$. By comparison, solar wind transports only about $3 \times 10^{-14} M_{\odot} y^{-1}$ (IMA2 p. 660).

For a white dwarf star of mass $0.85 M_{\odot}$ and radius $0.0095 R_{\odot}$, the potential gravitational energy U of an infalling test mass that is converted to KE (and ultimately to light and heat) is $1.71 \times 10^{13} J \text{ kg}^{-1}$, or **0.019%** of the rest energy (given by mc^2) of the test mass.

However, for a neutron star of mass $1.4 M_{\odot}$ and radius $10 \text{ km} = 1.4 \times 10^{-5} R_{\odot}$, the potential gravitational energy of an infalling test mass that is converted is $1.86 \times 10^{16} J \text{ kg}^{-1}$, or **21%** of the total mc^2 rest energy of the test mass. This is a prodigious amount of energy, 30 times what would be released from hydrogen fusion of the same mass of hydrogen. (IMA2 p. 661)

The observed luminosity of celestial X-ray sources are typically about 10^{30} W . If this power arises from infalling of mass to a neutron star, the mass transfer rate required is only about $10^{-9} M_{\odot} \text{ yr}^{-1}$.

Circumstellar Accretion Disks



Close binary system with red giant secondary star overflowing its Roche lobe across L1 into an accretion disk encircling a WD primary star.

The system is rotating counterclockwise, and this causes the inflowing gas from the secondary to miss the white dwarf. Instead, the rapid inflow first comes in contact with the periphery of the accretion disk at a bright **hot spot**.³⁹¹

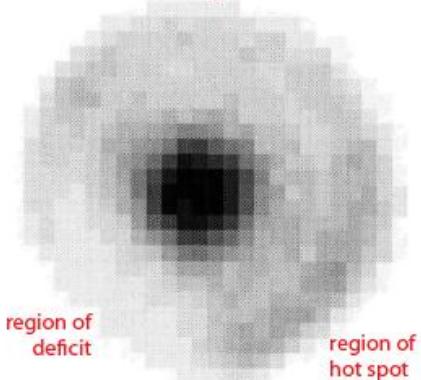
³⁹¹ <http://www.pas.rochester.edu/~afrank/A105/LectureXII/LectureXII.html>
Page 151 of 249

The orbital motion of a semidetached binary often prevents the direct infalling of the transported matter to the primary (companion), usually when the primary has a radius of <5% of the **binary separation** a (measured center to center). Instead the material goes into orbit about the primary to form a thin **accretion disk** in the plane of the orbits shared by the two stars.

Accretion Disk Hot Spot

The first region of contact, where infalling material collides with the accretion disk, can appear as a **hot spot** of enhanced intensity. This can be seen in the light curves of eclipsing binary systems and reconstructed into a circular intensity distribution. (Reconstructed negative image of LX Ser to right, slightly contrast enhanced by MCM)³⁹²

Reconstructed LX Serpentis Accretion Disk



Accretion Disk Viscosity and Heating

The gases and particles in the accretion disk collide and slow from **viscosity** (caused by still uncertain mechanisms). This process heats the disk, with greatest T attained in the innermost disk region, and light (or other EM) blackbody radiation is emitted. As the gas and particles slow, they spiral inward and eventually reach the surface of the primary. The migration period takes a few days to weeks. (IMA2 p. 666)

Accretion Disk Temperature, Luminosity, and Eddington Accretion Limit

The following formula best estimates the temperature of an accretion ring at a particular radius r (IMA2 p. 663):

$$T = \left(\frac{3GM\dot{M}}{8\pi\sigma R^3} \right)^{1/4} \left(\frac{R}{r} \right)^{3/4} (1 - \sqrt{R/r})^{1/4} = T_{disk} \left(\frac{R}{r} \right)^{3/4} (1 - \sqrt{R/r})^{1/4}$$

where the **characteristic temperature** T_{disk} is defined by

$$T_{disk} \equiv \left(\frac{3GM\dot{M}}{8\pi\sigma R^3} \right)^{1/4}$$

R and M are the radius and mass of the primary star, \dot{M} is the rate of matter infalling, σ = the Stefan-Boltzmann constant, and r is a radius measured from the center of the primary.

The **characteristic temperature** T_{disk} has a value approximately twice the maximum disk temperature (thus it is an overestimate), which occurs at $r \approx 1.36 R$. (IMA2 p. 707)

The luminosity of the accretion disk is given by (IMA2 p. 664):

$$L_{disk} = \frac{GMM\dot{M}}{2R}$$

The luminosity of the accretion disk shown here is 1/2 of the total luminosity arising from the infalling matter—the other half comes from impact at the star surface (rather than at the accretion disk).

The maximal rate of mass transfer to the accretion disk at which the Eddington accretion limit is reached is roughly given by³⁹³

$$\dot{M} = 4\pi c R / \kappa$$

³⁹² Rutten RGM, et al, “Reconstruction of the accretion disk in six cataclysmic variable stars”, *Astronomy and Astrophysics*, vol. 260, no. 1-2, p. 213-226, 07/1992.

³⁹³ Philipp Podsiadlowski, http://www-astro.physics.ox.ac.uk/~podsi/lec_c1_4_c.pdf

or $\approx 1.8 \times 10^{-8} M_{\odot} \text{ yr}^{-1}$ for a neutron star. Here, $\kappa \approx 0.034 \text{ m}^2 \text{ kg}^{-1}$ is Rosseland mean opacity, and R is the radius to the inner edge of the accretion disk. At this rate of mass deposition, the inner accretion disk is raised to heights reaching $2R$, where R is the radius of the NS (*IMA2* p. 693).

For the same typical white dwarf WD used above, with mass $0.85 M_{\odot}$ and radius $0.0095 R_{\odot}$, and for $\dot{M} = 1.6 \times 10^{-10} M_{\odot} \text{ yr}^{-1}$, the T_{\max} of the disk is $2.62 \times 10^4 \text{ K}$, the λ_{peak} is 111 nm (ultraviolet region), and the luminosity $L_{\text{disk}} = 8.55 \times 10^{25} \text{ W} = 0.22 L_{\odot}$.

For the same typical neutron star NS used above, with mass $1.4 M_{\odot}$ and radius $10 \text{ km} = 1.4 \times 10^{-5} R_{\odot}$, and for $\dot{M} = 1.6 \times 10^{-9} M_{\odot} \text{ yr}^{-1}$, the T_{\max} of the disk is $6.86 \times 10^6 \text{ K}$, the λ_{peak} is 0.423 nm (X-ray region), and the luminosity $L_{\text{disk}} > 2400 L_{\odot}$! (*IMA2* p. 665) Thus the accretion disks of neutron stars can be strong emitters of X-rays.

Accretion Disk Radial Extent

The accretion disk may not extend all the way to the surface of the primary star, due to magnetic field effects.

The outer limit of the accretion disk R_{disk} is estimated in terms the radius r_{circ} at which matter infalling from L_1 first settles into a circular orbit.

$$R_{\text{disk}} \approx 2r_{\text{circ}} = 2a \left[0.500 - 0.227 \log_{10} \frac{M_2}{M_1} \right]^4 \left(1 + \frac{M_2}{M_1} \right)$$

Here, a is again the binary separation (measured center to center).

For contact or semi-detached binaries, the transfer of mass from a star to its companion can also result in a significant transfer of angular momentum. Some of the mass that crosses L_1 into the primary's Roche lobe will carry a higher than average angular momentum, causing it to migrate to the outer edge of the disk. There it may be blown by the stellar wind, perhaps reaching the Roche lobe boundary and spilling into the "circumbinary" region or being blown from the binary system by the stellar wind, carrying away some of the system's angular momentum, just as the solar wind carries away angular momentum. (Gravitational radiation is another mechanism for carrying away angular momentum from relativistic binary systems, and is primarily responsible for loss of angular momentum in some short orbital period systems. (*IMA2* p. 669 and *here*³⁹⁴))

The gas that reaches the primary may add to the rotation of that star. Although tidally locked stars are not rotating with respect to each other, some binary systems have more complex stellar rotations and rotation axes. The text states however that stellar rotation angular momentum may be neglected. (*IMA2* p. 668)

A Survey of Interacting Binary Systems

As binary systems evolve, the ratio of their stellar masses and individual angular momenta change.

Effect of Mass Transfer

As mass is transferred, the orbital separation changes, according to this formula (*IMA2* p. 669):

$$\frac{1}{a} \frac{da}{dt} = 2\dot{M}_1 \left(\frac{M_1 - M_2}{M_1 M_2} \right)$$

Assuming as before that $M_1 > M_2$, mass transfer from the less massive star M_2 to the more massive one M_1 will make \dot{M}_1 positive, thus cause the RH to be positive, and the orbital separation a will increase. Mass transfer from the more massive one M_1 to the less massive star M_2 will make \dot{M}_1 negative, thus cause the RH to be negative, and the orbital separation will shrink. (Inferred from *IMA2* p. 669, explicit *here*.³⁹⁵)

As the orbital separation decreases, the angular frequency increases (orbital period decreases), according to (*IMA2* p. 669):

³⁹⁴ <http://arxiv.org/pdf/0811.0455>

³⁹⁵ <http://astronomy.nju.edu.cn/~lixn/XRB/chapter02.pdf>

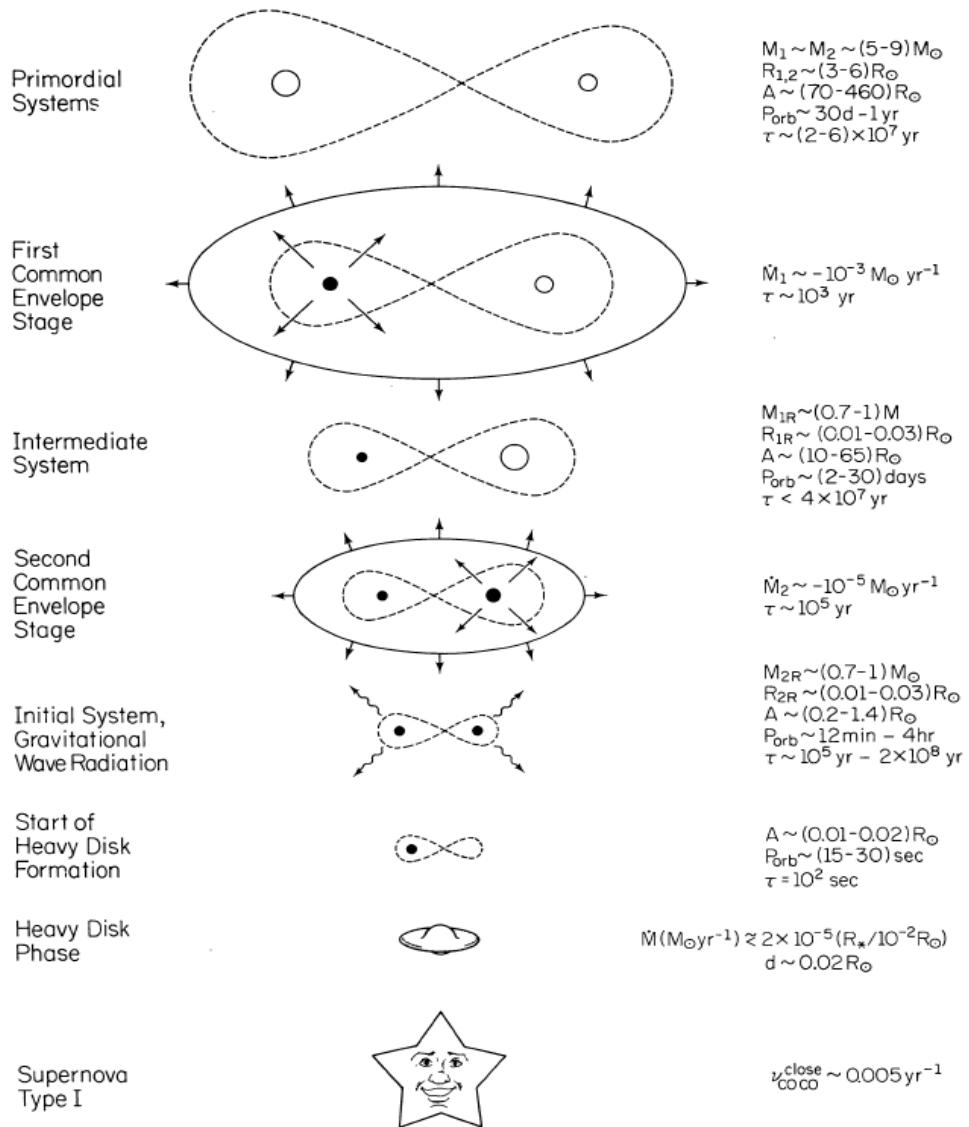
$$\frac{1}{\omega} \frac{d\omega}{dt} = -\frac{3}{2} \frac{1}{a} \frac{da}{dt}$$

Evolution of a Binary System

There are numerous possible evolutionary pathways for close binary systems (see *IMA2* p. 669–673).

The following diagram (*IMA2* p. 670-1 and *here*³⁹⁶) shows a representative sequence for a close binary system. It begins with two well-separated main sequence stars of similar mass and evolves to two carbon-oxygen (C-O) white dwarfs, one of which becomes a cataclysmic variable or Type Ia supernova. One star, M₁ (on left), becomes a red giant and overflows its Roche lobe to transfer mass to the less massive star M₂. This causes the stars to spiral closer together, producing a contact system with a common atmospheric envelope. In one scenario, these two cores might merge to form a *blue straggler*. But, as depicted here, the envelope may be ejected (similar to a planetary nebula). In this case, the system again emerges as a detached binary, with the primary M₁ now a white dwarf with exposed helium core, evolving to a C-O WD. Mass begins to transfer from M₂ to a new common envelope, and the stars spiral apart (per *IMA2* p. 670) or come closer (per the article and the diagram) as M₂ evolves also to a C-O WD. The larger less massive WD M₂ dissolves into a heavy disk that is accreted by the remaining WD M₁. Eventually, the remaining WD M₁ is pushed over the Chandrasekhar limit and explodes as a **Type Ia supernova**. $v_{CO CO}^{close}$ is the frequency (yr⁻¹) in the galaxy of SNe Ia from 2 C-O WD's by the "close channel"; τ is the duration of a stage; P_{orb} is the orbital period; A is the semi-major axis.

³⁹⁶ Iben I, Tutukov AV, "Supernovae of type I as end products of the evolution of binaries with components of moderate initial mass (M not greater than about 9 solar masses)", *Astrophysical Journal Supplement Series*, vol. 54, Feb. 1984, p. 335-372, <http://adsabs.harvard.edu/abs/1984ApJS...54..335>



Types of Interacting Binary Systems

"We define an active binary as a binary stellar system in which the activity level is significantly enhanced as a consequence of it being a binary with respect to the level which is typical for single stars of the same mass and evolutionary state."³⁹⁷

Interacting Binary Systems include (*IMA2* p. 672-3 and *here*³⁹⁸):

- Algols — Algol systems consist of two ‘normal’ main sequence or subgiant stars in a semi-detached system. They contribute mass to the ISM. Algol is an eclipsing binary consisting of a B8 and K2 star.³⁹⁹
- W Serpens Stars — Active Algol systems.
- RS Canum Venaticorum (RS CVn) Stars⁴⁰⁰ — Chromospherically active close binaries of F (cool) stars and later spectral type components. They have large stellar spots which may cause luminosity to vary.
- BY Draconis Stars — Similar to RS Canum Venaticorum systems. “The variability of BY Draconis is caused by activity in the stellar photosphere called starspots, which are comparable to sunspots on the Sun, in combination with rapid rotation that changes the viewing angle of the activity relative to the observer. This

³⁹⁷ <http://articles.adsabs.harvard.edu/full/1995A%26A...298..482F>

³⁹⁸ www.astro.caltech.edu/~george/ay20/Ay20-Lec11x.pdf

³⁹⁹ <http://csep10.phys.utk.edu/astr162/lect/binaries/algol.html>

⁴⁰⁰ http://en.wikipedia.org/wiki/RS_Canum_Venaticorum_variable

variation has an average periodicity of 3.8285 days, but the brightness also changes over the course of several years—depending on the level of surface activity. Most observers believe that the primary star (A) is responsible for the variability as the secondary produces only a third of the total luminosity from the system. However, the spots may occur on both stars. Unlike the Sun, these spots may occur in the polar regions of the stars.”⁴⁰¹

- Cataclysmic Variables (CV) and Nova-like binaries⁴⁰²—Have short orbital periods and contain a WD primary component together with a cool M-type secondary that fills its Roche lobes. These irregularly and dramatically increase in brightness as mass transfers lead to bursts of surface nuclear fusion.
- W Ursae Majoris (W UMa) Contact Systems⁴⁰³ — Short orbital period (0.2 to 0.8 day) contact binaries exhibiting very high levels of magnetic activity. Magnetic braking may cause these stars to coalesce.
- X-ray Binaries⁴⁰⁴ — Binaries with a neutron star or, more rarely, a black hole primary. Accretion onto the degenerate primary produces strong x-ray emissions. These provide evidence for the existence of black holes.
- Zeta Aurigae (ζ Aur) Systems⁴⁰⁵ — Long orbital period binaries comprising interacting type G or K supergiant and hot (~type B) companion. Although not originally interacting systems, they become so when the more massive star evolves to become a supergiant.
- VV Cephei Systems⁴⁰⁶ — Similar to Zeta Aurigae binaries with long orbital period, except a type M supergiant replaces the G or K type supergiant.
- Symbiotic Binaries — Long orbital period (~200 to 1500 days) interacting binaries consisting of a cool type M giant (sometimes a pulsating Mira-type variable) and a hot accreting companion such as a white dwarf, sub-dwarf, or low-mass main sequence star. “The hot star ionizes some of the material shed by the giant, and this leads to the symbiotic's characteristic composite spectrum of nebular emission lines superimposed on molecular absorption bands. Symbiotic binaries are valuable laboratories for the study of cool stellar winds, ionized nebulae, wind-fed accretion processes and wind-wind collision dynamics.”⁴⁰⁷ Where the cool star fills its Roche lobe, the system becomes a symbiotic Algol.
- Barium and S-Star Binaries — Long period binaries in which the originally more massive component evolved into a white dwarf, transferring some of its nuclear-processed gas to a giant type K or M companion. (S spectral type stars are intermediate between M and C stars.) “Barium stars are spectral class G to K giants, whose spectra indicate an overabundance of s-process elements by the presence of singly ionized barium, Ba II, at λ 455.4nm. Barium stars also show enhanced spectral features of carbon, the bands of the molecules CH, CN and C₂. ”⁴⁰⁸
- Post-Common Envelope Binaries — Usually consisting of a hot white dwarf or subdwarf, and a cooler secondary, which have passed through the common envelope phase. The binary nuclei of planetary nebulae are examples.

White Dwarfs in Semidetached Binaries

If the WD is the primary, these systems can exhibit dwarf novae, classical novae, or supernovae Type Ia.

⁴⁰¹ http://en.wikipedia.org/wiki/BY_Draconis

⁴⁰² Cataclysmic Variable Stars:

• Szkody et al, “Cataclysmic Variables From Sloan Digital Sky Survey. IV. The Fourth Year (2003)”, 2007, <http://iopscience.iop.org/1538-3881/129/5/2386/.../204549.text.html>

• http://en.wikipedia.org/wiki/Cataclysmic_variable_star

⁴⁰³ http://en.wikipedia.org/wiki/W_Ursae_Majoris

⁴⁰⁴ http://en.wikipedia.org/wiki/X-ray_binary

⁴⁰⁵ <http://www.hposoft.com/EAur09/zeta%20Aurigae/zeta.html>

⁴⁰⁶ http://en.wikipedia.org/wiki/VV_Cephei

⁴⁰⁷ <http://www.astro.ex.ac.uk/people/th2/symbo.html>

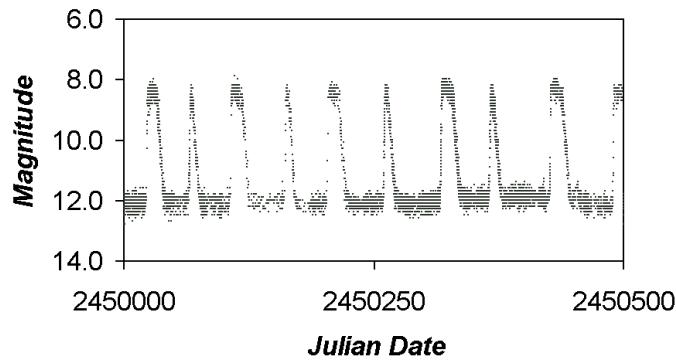
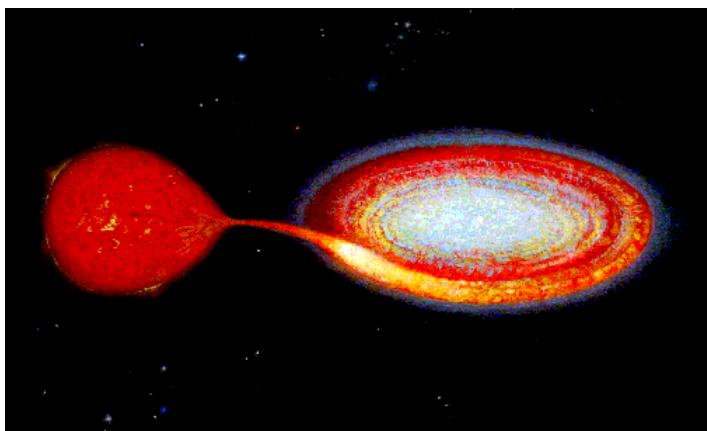
⁴⁰⁸ http://en.wikipedia.org/wiki/Barium_star

Cataclysmic Variables (CV)

This is a favorite research interest for Professor Szkody.⁴⁰⁹ There are over 1600 known CV systems.⁴¹⁰ These are subdivided into dwarf novae, classical novae, polars, and other less common types. CVs are “binary star systems that have a white dwarf and a normal star companion. They are typically small—the entire binary system is usually the size of the Earth-Moon system—with an orbital period of 1 to 10 hours. The white dwarf is often referred to as the ‘**primary**’ star, and the normal star as the ‘**companion**’ or the ‘**secondary**’ star. The companion star, a star that is ‘normal’, like our Sun, loses material onto the white dwarf via accretion. Since the white dwarf is very dense, the gravitational potential energy is enormous, and some of it is converted into X-rays during the accretion process. There are probably more than a million of these cataclysmic variables in the [MW] galaxy, but only those near our Sun (several hundreds) have been studied in X-rays so far. This is because CVs are fairly faint in X-rays; they are just above the coronal X-ray sources and far below the X-ray binaries in terms of how powerful their X-ray emissions are.”⁴¹¹

Dwarf Novae

Dwarf novae have several subtypes, including **Z Camelopardalis** stars, **SU Ursae Majoris** stars, **SS Cygni** stars . “[Their] outbursts result from temporary increases in the rate of accretion onto the white dwarf, caused by the additional material accreted onto the surface. This material must go through a violent transition region called the ‘boundary layer’, which lies just above the surface of the white dwarf. Dwarf novae outbursts are smaller in amplitude and higher in frequency than classical novae. The variable star U Geminorum, or ‘U Gem,’ is the prototype of dwarf novae. The brightness in the visible light of U Gem increases by a hundredfold every 120 days or so, and returns to the original level after a week or two. Optical astronomers have also recognized ‘**recurrent novae**,’ which are eruptive behaviors that fall between the definitions of classical and dwarf novae, and ‘nova-like systems,’ which are stars that have similar spectra to other types of CVs in the visual light, but have not been seen to erupt.”⁴¹² *IMA2* p. 675 states that the brightening takes place on the accretion disk, rather than at the WD surface, and that the outbursts are caused by a sudden increase in the rate at which mass flows down through the accretion disk. The T_{\max} of the accretion disk of Z Camelopardalis during an outburst is 4.4×10^4 K, the mass transfer rate is estimated at $1.3 \times 10^{-9} M_{\odot} \text{ yr}^{-1}$ (but lasting only days), and the total luminosity is $6.8 \times 10^{26} \text{ W}$, which is about twice L_{\odot} . The durations of dwarf novae outbursts are typically 5 to 20 days, separated by 30 to 300 days of quiescence. The mass transfer rate increases by a factor of about 10 to 100, similar to the brightness. (A 100x increase in brightness is equivalent to a 5 magnitudes increase.) The mechanism of the irregular outbursts is unclear, perhaps disk instabilities. (see *IMA2* p. 676-8 for hypotheses) The graph below shows “A portion of the SS Cygni light curve from the AAVSO International Database displaying interchanging bouts of wide and narrow outbursts.”⁴¹³



⁴⁰⁹ <http://www.astro.washington.edu/users/szkody>

⁴¹⁰ http://en.wikipedia.org/wiki/Cataclysmic_variable_star

⁴¹¹ http://imagine.gsfc.nasa.gov/docs/science/know_12/cataclysmic_variables.html

⁴¹² ibid.

⁴¹³ http://www.aavso.org/vsots_ssccy

(Left) CV dwarf nova binary system Z Camelopardalis: Painting by Dale W. Bryner⁴¹⁴ shows Roche lobe overflow and hot spot at region of impact.

(Right) Light curve segment of dwarf nova system SS Cygni, showing alternating wide and narrow outbursts.⁴¹⁵

Polars and Miscellaneous CVs

Polars are also CVs. The **AM Herculis** stars are stars where the magnetic field of the white dwarf has locked the binary system into nearly synchronous rotation. Matter streams straight down an accretion column fairly constantly onto the white dwarf rather than forming a disc. They have fields of about 2000 T and emit X-ray and UV light. The **DQ Herculis** types are also called **intermediate polars**. They have a weaker magnetic field than AM Herculis stars, namely <1000 T. They form an accretion disk but do not exhibit synchronous rotation. The visible light they emit is from *nonrelativistic* electrons spiraling along the magnetic field lines of the accretion column. This light is called *cyclotron* (not synchrotron) radiation. It alternates between circular and linear polarization, a phenomenon giving these stars the name ‘polars’. (*IMA2* p 685 and *here*⁴¹⁶)

Additional CV types include **VY Sculptoris**, **AM Canum Venaticorum**, and **SW Sextantis**.

Novae (Classical Nova)

The first nova witnessed was CK Vulpeculae in 1670. About 30 to 60 probably occur per year in the MW, but only about 10 per year are identified due to obscuration by dust, etc. About 30 are seen annually in Andromeda Galaxy and other nearby galaxies.⁴¹⁷ Classical novae have higher accretion rates than dwarf novae. They have very large outbursts, of 7 to 20 magnitudes, caused by a runaway thermonuclear fusion of material accreted onto the white dwarf. “The amplitude of the outburst is the largest among CVs. Classical nova outbursts are caused by sudden nuclear fusion of hydrogen-rich material on the surface of the white dwarf...”⁴¹⁸ Hydrogen fusion in a WD is possible only when fresh fusible fuel is accreted onto its surface.

In a nova, the primary WD averages $0.86 M_{\odot}$, the secondary is usually a MS star of type G or later and less massive than the WD. Orbital periods tend to be 78 to 720 minutes. During the outburst, the accretion disk is more optically thick, suggesting more material is flowing through it. Doppler shifted emission lines appear during the quiescent phase... (see *IMA2* p. 673)

The peak temperature attained in the localized fusion process is as high as 10^8 K, and radiation pressure may cause some of the accreted material to be blown away (the **hydrodynamic ejection phase**). A prolonged equilibrium **hydrostatic burning phase** later sets in, which includes some CNO fusion. The hydrostatic burning phase eventually ends and cooling begins, finally returning to the quiescent configuration, at which point accretion resumes. However, it may take 10^4 to 10^5 years to build up a surface layer of sufficient size ($10^{-4} M_{\odot}$) to cause a recurrent thermonuclear fusion reaction and outburst. The mass transfer rate is estimated at $9.0 \times 10^{-10} M_{\odot} \text{ yr}^{-1}$ (but lasts much longer than dwarf novae). (*IMA2* p. 682, 684) Several recurrent novae have been identified.

⁴¹⁴ <http://physics.weber.edu/carroll/>

⁴¹⁵ http://www.aavso.org/vsots_sscyg

⁴¹⁶ http://en.wikipedia.org/wiki/Cataclysmic_variable_star

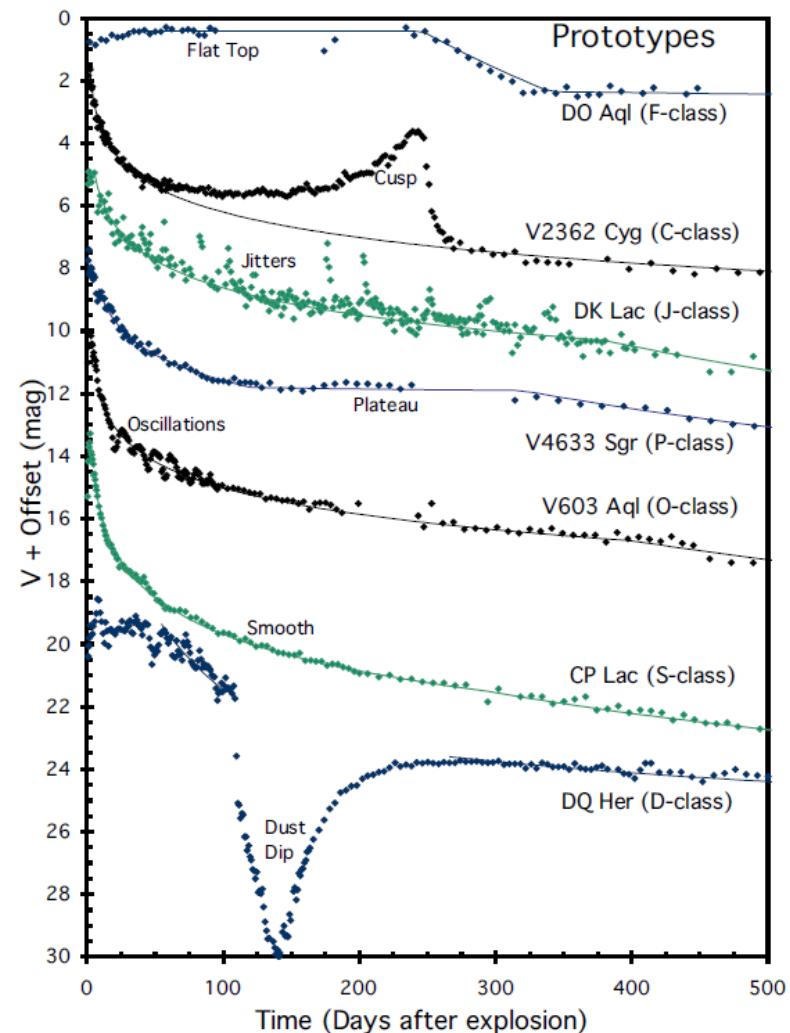
⁴¹⁷ <http://en.wikipedia.org/wiki/Nova>

⁴¹⁸ http://imagine.gsfc.nasa.gov/docs/science/know_12/cataclysmic_variables.html

The rise to peak luminosity for a nova is rapid, and takes only a few days. At peak brightness, these can have luminosity of $10^5 L_\odot$. During an outburst, V-band luminosity falls off faster than V+UV (bolometric) luminosity. (IMA2 p. 682) The decline from the peak by 2 magnitudes takes a few weeks for a “**fast nova**” NA and up to 100 days for a “**slow nova**” NB. Light output can be very irregular, however. (IMA2 p. 680 and here⁴¹⁹) Strope et al 2010 describe a variety of light curves for novae, designated as S, P, D, C, O, F, and J.⁴²⁰ For example, the D class light curves stands for **Dust Dip**: “The defining feature of D class light curves is a dust dip, which starts suddenly after the initial decline. Dust dips appear as steep drops in the brightness, although the depth of this drop can vary. The dust dips then reach some minimum brightness, followed by a recovery in brightness to a fainter level than the start of the dip, followed by a slow decline. The classic D light curve is that of DQ Her [DQ Herculis, a slow nova]. However, dust dips occur with a very wide range of depths and durations. The dust dip of V445 Pup is over ten magnitudes in depth, taking the star to over four magnitudes fainter than its normal quiescent level, with the dip remaining nearly this deep at least until 2009... At the other extreme are shallow dips...” The seven light curve types are depicted in the graphs to the right. The words for which the letters S, P, D, C, O, F, and J are the first initial are spelled out in the graph labels. “Each nova has its V-band light curve with an arbitrary offset added so that the light curves are well separated, with the shape being retained. Curved line segments are superposed on the binned and averaged magnitudes (diamonds) to represent the fitted broken power law models.”

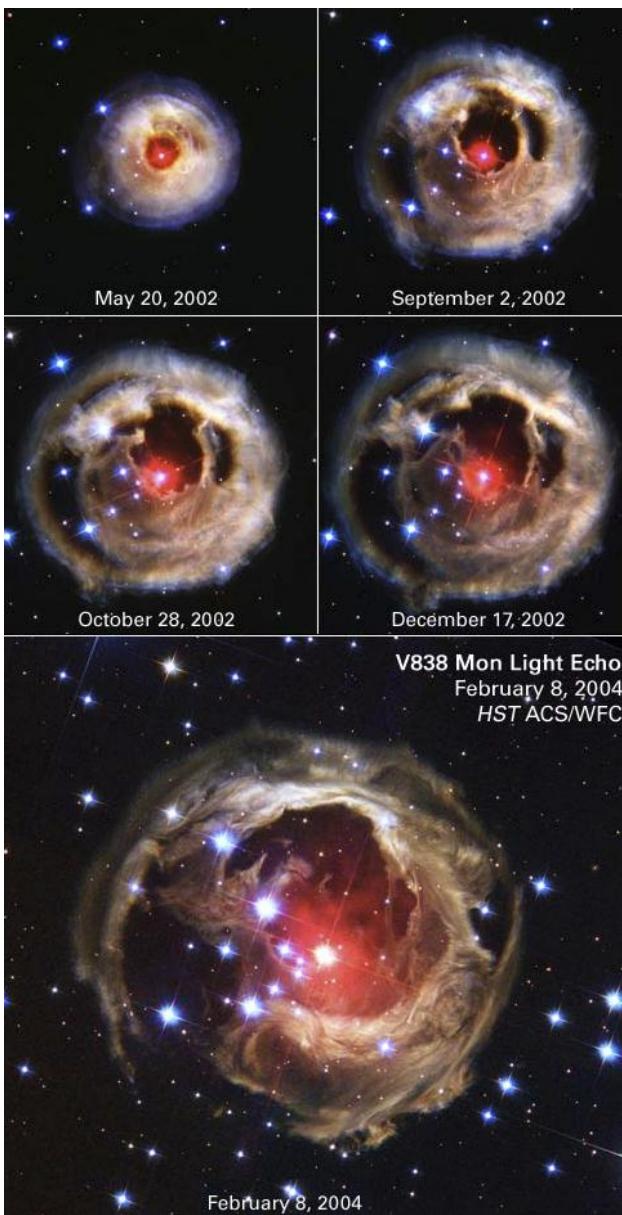
The ejected gases pass through several stages. The **fireball expansion phase** resembles an A or F supergiant, and radiates at 6000–10000K. (IMA2 p. 684) The initially optically dense fireball (including the time of peak luminosity) becomes less and less dense as it continues to expand, eventually becoming transparent in the

optically thin phase. A limiting fireball photospheric radius of about 5×10^{10} m is attained. (I have some uncertainty about this statement.) After a few months, the carbon in the ejecta begins to coalesce as graphite grains and dust forms, beginning the **dust formation stage**. The shell becomes optically thicker, and in the visible light band, the central star may be hidden. However, bolometric luminosity remains high when the IR emitted by the dust at about 900 K is included. The expanding shell may remain visible for many years after the hydrostatic burning phase has ended. (IMA2 p. 682, 685)



⁴¹⁹ http://en.wikipedia.org/wiki/Cataclysmic_variable_star

⁴²⁰ Strope RJ, et al, “Catalog of 93 Nova Light Curves: Classification and Properties”, <http://arxiv.org/pdf/1004.3698v1>



Evolution of an atypical nova-like red variable star V838 Monocerotis (V838 Mon) and surrounding *light echo* (HST)

Regarding this sequence of images, this event was initially thought to be a nova. However, the Hubble Site now states, “The illumination of interstellar dust comes from the red supergiant star at the middle of the image, which gave off a flashbulb-like pulse of light two years ago. V838 Mon is located about 20,000 light-years away from Earth in the direction of the constellation Monoceros, placing the star at the outer edge of our Milky Way galaxy... Astronomers do not fully understand the star's outburst. It was somewhat similar to that of a nova, a more common stellar outburst... [But unlike a nova,] V838 Monocerotis did not expel its outer layers. Instead, it grew enormously in size. Its surface temperature dropped to temperatures that were not much hotter than a light bulb. This behavior of ballooning to an enormous size, but not losing its outer layers, is very unusual and completely unlike an ordinary nova explosion... The outburst may represent a transitory stage in a star's evolution that is rarely seen. The star has some similarities to highly unstable aging stars called eruptive variables, which suddenly and unpredictably increase in brightness.⁴²¹ Although not a typical nova, I could not pass up including these remarkably elegant HST images.

⁴²¹ <http://hubblesite.org/newscenter/archive/releases/2004/10/> image rearranged by MCM
Page 160 of 249

Type Ia Supernovae

Unlike Type Ib, Ic, and II SNe, Type Ia SN vary relatively little in luminosity and rate of decline, etc., and generally they vary in predictable ways that can be adjusted for. Aside from their intrinsic interest, these serve an extremely important role as standard candles for the measurement of distances to the galaxies in which they reside, and therefore of extragalactic and cosmological distances. They have provided critical evidence that the rate of expansion of the universe has been accelerating ($\ddot{a}(t)$, the second derivative of the cosmic scale factor, is >0). This role, and the methods of analysis of light curves, luminosity, etc. are discussed in Chapter 27.

The typical peak luminosity expressed as absolute magnitude is $\langle M_B \rangle \cong \langle M_V \rangle \cong -19.3 \pm 0.3$ (IMA2 p. 686 and erratum). In the scaled magnitude light curve (chapter 27), the time to peak is less than 20 days, and averages about 19 days.⁴²²

They lack H lines (like all Type I SNe), but exhibit strong Si II lines (including 615 nm), along with lines of neutral and ionized O, Mg, S, Ca, and Fe. (IMA2 p. 687).

The mass ejected is reflected in P Cygni profiles (see Chapter 12) and the rates of expansion are shown by blueshift to be as much as 10^7 m s^{-1} ($\sim 0.1c$).

As discussed also in Chapter 16 and 27, the model of the Type Ia SN starts with a WD in a binary systems that is approaching the Chandrasekhar Limit for White Dwarfs of about $1.4 M_{\odot}$. The companion overflows its Roche lobe and contributes additional mass to the WD until the limit is reached or crossed, at which point a SN results.

A “**double degenerate**” model⁴²³ depicts two WDs in a binary orbiting system. As they orbit, they give off gravitational waves that remove energy and angular momentum, causing the stars to spiral in toward each other. When sufficiently close, tidal forces break up the less massive larger WD in a few orbits and C-O rich mass is dumped onto and merges with the other WD. Nuclear reactions begin deep in the interior and eventually destroy the WD in a Type Ia SN (IMA2 p. 688)

An alternative and partially more favored model, however, is the “**single degenerate**” model. A WD has a normal star as a companion, the latter overflows its Roche lobe, the WD attains the Chandrasekhar Limit, and a SN Type Ia results, destroying at least the WD. This model would seem to lead to a more uniform mass just at the Chandrasekhar Limit at the time of the SN. A helium flash may start the explosion, and eventually the deeper C-O ignites in a deflagration (subsonic) or detonation (supersonic explosion), with internal temperatures of billions of K.

In any event, the SN Type Ia ultimately results from the eruption of a C-O WD at or near the Chandrasekhar limit of $1.4 M_{\odot}$.⁴²⁴ The energy released is $1-2 \times 10^{44} \text{ J}$.⁴²⁵ The mass ejected is reflected in P Cygni profiles (see Chapter 12) and the rates of expansion are shown by blueshift to be as much as 10^7 m s^{-1} . (IMA2 p. 689) Another sources states that the star releases a shock wave in which matter is typically ejected at speeds on the order of 5 to $20 \times 10^6 \text{ m s}^{-1}$, thus as high as $0.07c$.⁴²⁶

Neutron Stars and Black Holes in Binaries

After a core-collapse SN in one of the stars in a binary system (or an isolated NS capture), a NS or black hole BH may be left in a binary system in which the compact degenerate star is orbiting the companion star. When the companion overflows its Roche lobe, the gas infalling onto the compact star or its accretion disk is associated with copious release of X-rays.

Supernova Explosion Impact on Binary Systems

After a core-collapse SN in a star M_1 in a binary system, leaving a compact star with remnant mass M_R , and neglecting the direct blast effects on the companion star:

⁴²² <http://en.wikipedia.org/wiki/Supernova>

⁴²³ See the Iben and Tutkov 1984 diagram earlier in this chapter

⁴²⁴ <http://www.cfa.harvard.edu/news/2012/pr201214.html>

⁴²⁵ http://en.wikipedia.org/wiki/Type_Ia_supernova

⁴²⁶ http://en.wikipedia.org/wiki/Type_Ia_supernova

If $\frac{M_R}{M_1+M_2} \leq 1/2$ (the ratio of post-SN M₁ mass to total initial binary mass $\leq 1/2$),
the system is *unbound*

If $\frac{M_R}{M_1+M_2} > 1/2$ (the ratio of post-SN M₁ mass to total initial binary mass $> 1/2$),
the system becomes a NS or BH *bound* to the companion star M₂.

IMA2 p. 690 states that for M₂ \gg M₁, the latter case is likely (the description appears inconsistent however).

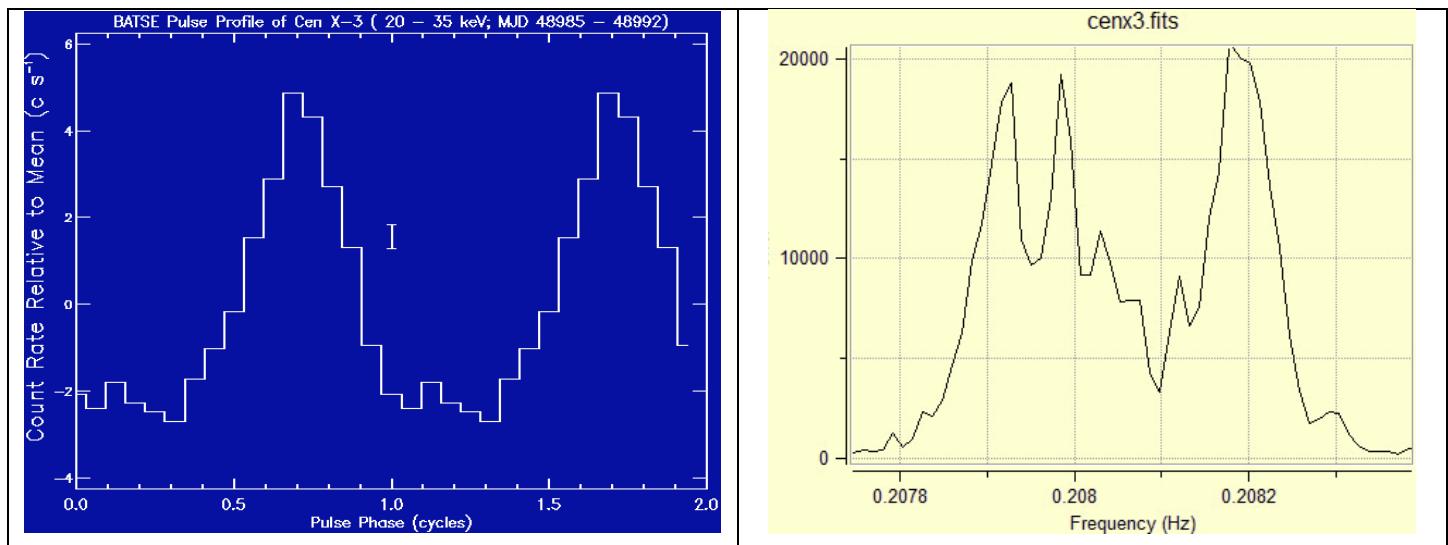
Capture of an Isolated Neutron Star

It is possible for a free NS resulting from a core-collapse SN to be subsequently captured by close approach to another star. If the excess kinetic energy can be dissipated through tidal effects on the nondegenerate star, a **tidal capture** results and the resulting close binary orbit will have an orbital period of hours to days. This is most likely to happen in dense star regions, such as the center of globular clusters, and can be expected to happen only 10 times in 10¹⁰ years. Such a close binary X-ray system lasts only about 10⁹ y, so X-ray binaries created by this interaction are expected to be rare. Additional hypothesized capture mechanisms include 3-body star interactions (in which the lightest star is ejected to carry away excess energy while the NS is captured) and penetration of a NS into a giant star, the latter leaving a WD and NS binary with an orbital period of about 10 minutes.

Binary X-ray Pulsars

These are also called X-ray pulsars. The first discovered, in 1962, was Scorpius X-1 (Sco X-1), which because of its proximity (2800 pc) is the strongest X-ray source in the sky other than the Sun. Its X-ray luminosity is about 2.3×10^{31} W. It is a NS binary accreting matter at its Eddington accretion limit from a low-mass donor companion star, and has an orbital period of around 18.9 hours.⁴²⁷

Centaurus X-3 (Cen X-3) is another strong binary X-ray pulsar, the first binary pulsar to be discovered in X-rays in 1971. The pulsar has a NS spin period of only 4.8 s and a binary orbital period of 2.09 d. This was the first X-ray pulsar discovered. It consists of a 1.2 M_⊙ NS that is regularly eclipsed by a massive 20.5 M_⊙ O-type supergiant companion which is feeding mass to the NS accretion disk.



Left graph: Typical X-ray emission profile of X-ray pulsar Cen X-3
for two full NS rotation periods of 4.8 s each⁴²⁸

⁴²⁷ http://en.wikipedia.org/wiki/Scorpius_X-1

⁴²⁸ <http://www.batse.msfc.nasa.gov/batse/pulsar/data/sources/cenx3.html> image modified MCM

Right graph: Power spectrum showing variation in pulsar frequency resulting from cyclical Doppler shifts as pulsar orbits its companion.⁴²⁹

X-ray pulsars have X-ray luminosity up to the Eddington limit of 10^{31} W and a T_{eff} of 2×10^7 K, giving a peak Wien emission $\lambda = 0.15$ nm, $E = 8.27$ keV, therefore in the X-ray band. (IMA2 p. 691).

X-ray pulsars also emit in radio frequencies, but this is not a prominent feature. Radio pulsars are mostly single rapidly rotating NS's, whereas X-ray pulsars are binary systems, mostly accretion-powered and consisting of a NS and an overflowing companion, and having slower and often more irregular rotation periods.⁴³⁰

X-ray pulsars like other NS's often have strong magnetic fields. These interact with the inflowing matter on the accretion disk to divert mass away from the disk plane toward the poles. The disruption of radial inflow and diversion takes place at a radius of $\sim 0.5r_A$, where the Alfvén radius r_A is given by:

$$r_A = \left(\frac{8\pi^2 B_S^4 R^{12}}{\mu_0^2 G M \dot{M}^2} \right)^{1/7}$$

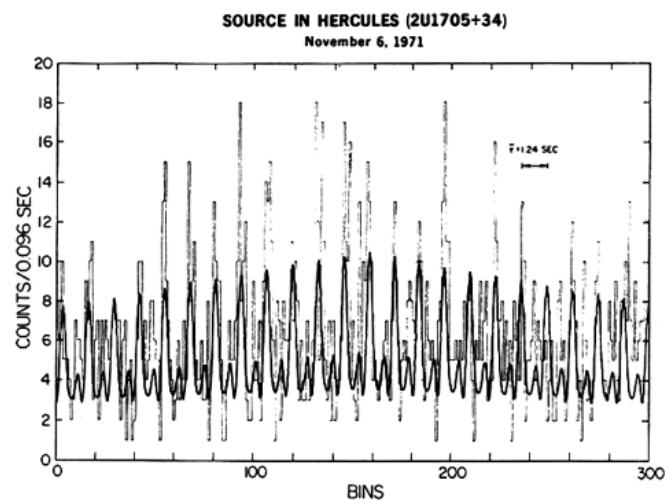
and where B_S = magnetic field at the surface, R is the NS surface radius. The channeled accretion, in a typical strong magnetic surface field $B_S = 10^8$ T, yields an Alfvén radius $r_A = 3.1 \times 10^6$ m and therefore a disruption radius of about 1.5×10^6 m. This is 150x the 10^4 m NS radius R but still much less than the radius at the outer disk margin ($2r_{\text{circ}}$) previously estimated. The accretion luminosity of the inner disk rises to the Eddington limit, and the inner disk expands to up to $2R$ in height out of the disk plane. The process creates a channel in which the particles follow the curving field lines and are funneled into the polar regions rather than proceeding to the NS equator. In these polar regions, the infalling matter creates hot spots on the circumpolar stellar surfaces, leading to a pulsating beam as the star rotates with the magnetic axis differing from the rotational axis. (If the magnetic field is especially strong, the matter may never actually reach the NS surface.) However, in contrast to a quasar, the X-rays are emitted over a large solid angle, especially in older NS's where the magnetic field has become weaker—a sufficiently wide spread may suppress the periodic variation entirely. See below about X-ray bursters.

Eclipsing Binary X-ray Systems

A NS with magnetic axis differing from rotational axis may be an eclipsing binary X-ray pulsar. A variety of mechanisms for the eclipses are proposed. IMA2 p. 695 states that there are about 20 known binary X-ray pulsars. These rotate too fast to be WD's. The frequency of x-ray pulsar rotation increases with time (spin period decreases), due to magnetic torque effects, according to a complicated formula (IMA2 p. 695).

The NS of **Hercules X-1** binary system rotates with spin period 1.24 s and generates short-term pulses with this period. Furthermore, it exhibits eclipses every 1.7 d with the period of the binary orbit. It also undergoes an almost periodic longer cycle consisting of 2 bright and 2 faint phases every 35 days. Excursions of 6 days in both directions have been observed about these 35-d cycles.

In the diagram to right, the short-term pulsations resulting from NS rotation are shown as counts per bin of 0.096 s over about 30 s.⁴³¹



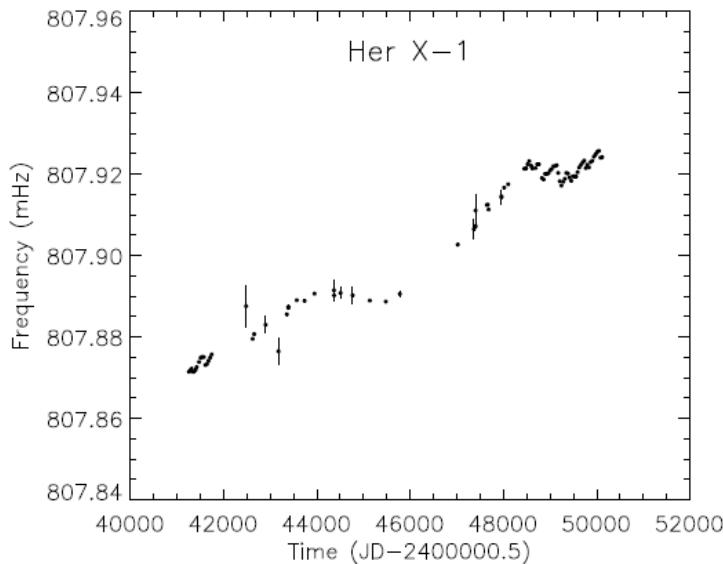
⁴²⁹ http://chandra-ed.harvard.edu/clocks/clocks_A3.html image modified MCM

⁴³⁰ http://en.wikipedia.org/wiki/X-ray_pulsar

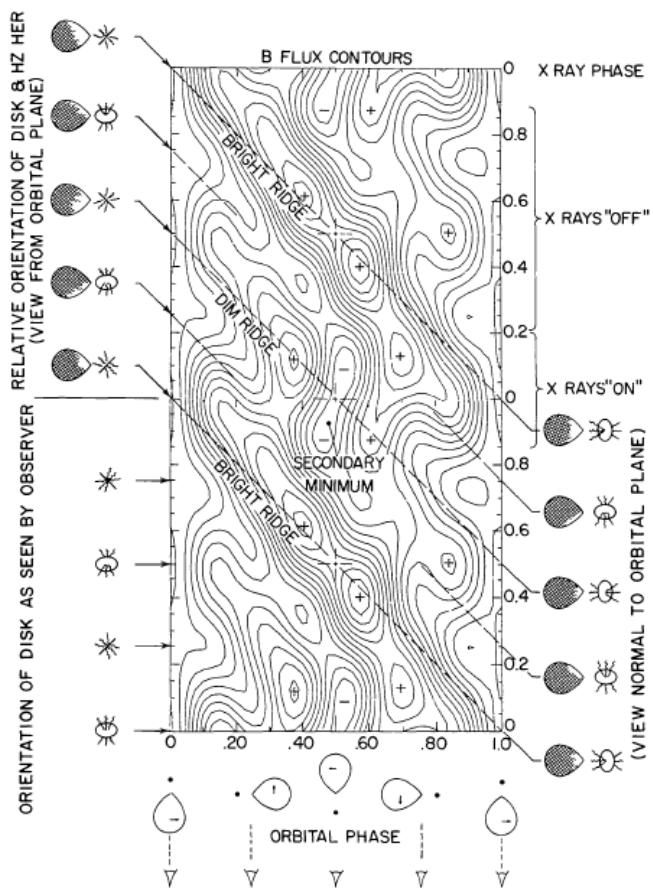
⁴³¹ http://heasarc.gsfc.nasa.gov/docs/objects/binaries/herx1_spin.html

In the optical flux diagram to the right,⁴³² the 1.7 day orbital period phase is depicted on the horizontal axis and the longer 35 day cycle is on the vertical axis ("X-ray phase"). Optical flux is a function of these two cycles (plus of course the 1.24 s pulsar rotation.) "The backward-marching intensity ridges are evidence for counter-precessing structure in the illumination of the heated companion star [HZ Her].” In other words, it is believed that the NS is precessing and that a tilted, twisted accretion disk in retrograde (backward marching) precession modulates the X-ray illumination and reprocessed X-ray flux of companion HZ Her (and therefore of the Earth), and this can explain the 35-day intensity ridge cycle.

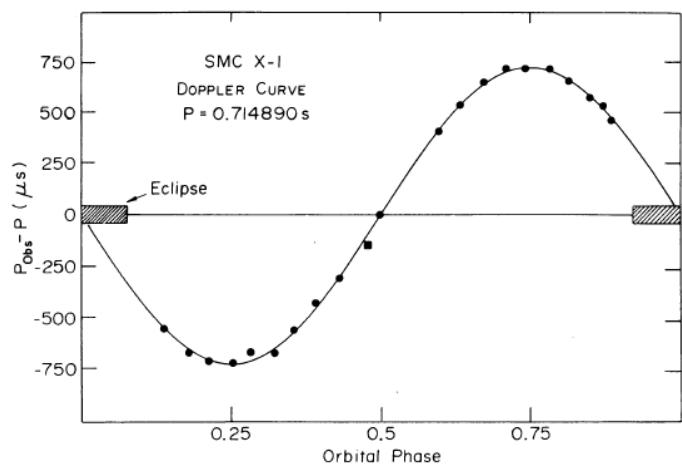
The diagram below⁴³³ shows the gradual increase in rotational frequency of the pulsar over an interval of about 10,000 days. Analysis shows that this rate of increase is much more compelling evidence for a NS as the accreting star rather than a WD.



unclear which) causes a cyclic variation in the measured pulse period (arrival time) as a function of orbital phase. In the graph⁴³⁴ to the right, the time delay in μs between observed period and constant period is plotted against orbital phase for the SMC X-1/Sanduleak 160 binary system. This system has a 3.89 day orbital period, and a pulse period = 0.715 s. It is said to be the most luminous known X-ray pulsar.



As the NS orbits, it changes distance from Earth in a sinusoidal pattern and its relative velocity is also varying sinusoidally. One or both effects (I am



⁴³² Friedhorsky WC, Holt SS, “Long-term cycles in cosmic X-ray sources”, *Space Science Reviews*, vol. 45, no. 3-4, 1987, p. 291-348. <http://dx.doi.org/10.1007/BF00171997>

⁴³³ Bildstein L, et al, “Observations of Accreting Pulsars”, *The Astrophysical Journal Supplement Series*, 113:367-408, 1997 December, <http://dx.doi.org/10.1086/313060>

⁴³⁴ Primini F, et al, “Orbital elements and masses for the SMC X-1/Sanduleak 160 binary system”, *Astrophysical Journal*, vol. 210, Dec. 1, 1976, pt. 2, p. L71-L74. <http://dx.doi.org/10.1086/182306>

Magnetic X-ray Bursters (XRB)

If the magnetic field is sufficiently low, the H- and He-rich matter can with time accumulate sufficiently near the surface to lead to release of bursts of thermonuclear energy lasting apparently only a few seconds, and recurring in hours to days. The fusion of H begins just below the surface, and He just below that. The He burning is explosively rapid, releasing $\sim 10^{32}$ J in a few seconds, with transient surface T as high as 3×10^7 K. An optical flash from the accretion disk may follow the X-ray burst by a few seconds. The rate of cooling suggests blackbody radiation from an object of radius 10 km, thus a NS. Such X-ray pulsars are called **X-ray bursters XRBs**. (IMA2 p. 691–694, 696, and *here*⁴³⁵) “Observationally, X-ray bursts are put into two distinct categories, labeled Type I and Type II. A Type I X-ray burst has a sharp rise followed by a slow and gradual decline of the luminosity profile. A Type II X-ray burst exhibits a quick pulse shape and may have many fast bursts separated by minutes. However, only from two sources have Type II X-ray bursts been observed, and so most X-ray bursts are assumed to be of Type I.”⁴³⁶ IMA2 p. 697 says there are about 50 known, but a list of about 100 Type-I X-ray bursters can be seen *here*.⁴³⁷ “X-ray bursts are seen coming from Globular Clusters, which are in the halo of our galaxy [the MW], and from sources along the plane of our galaxy”⁴³⁸ (whereas gamma ray bursts are distributed isotropically and are therefore mostly extragalactic).

Low-Mass X-ray Binaries (LMXB) and Massive X-ray Binaries (MXRB)

Mass determination in binary systems can be difficult, and in some cases only the ratio of the 2 stellar masses is known (for this, inclination is irrelevant, Chapter 7) or a lower limit on the mass can be deduced (uncertain due to unknown inclination). The X-ray Binaries are subdivided on the basis of the mass of the donor companion M_2 (not the compact star M_1 or total mass $M_1 + M_2$):

- **Low-Mass X-ray Binaries LMXBs:** These are semi-detached binaries consisting of either a neutron star or a black hole primary, and a low-mass secondary which is filling its critical Roche lobe. The companion star is usually a late spectral type star with traditionally $M_2 \lesssim 1.5 M_\odot$ (sometimes stated as $< 1 M_\odot$ or $< 2 M_\odot$), but can also be a WD or Spectral Type A star (Liu 2001). Typically, $L_{\text{optical}}/L_{\text{X-ray}} < 0.1$, and often undetectable in visible light. They are powered by mass transfer via overflow at L_1 . They are seen in older stellar populations—about 25% are in globular clusters, where the probability of NS capture is higher. Examples include Scorpius X-1 (Sco X-1), which has a $1.4 M_\odot$ NS accreting at its Eddington Limit,⁴³⁹ and A0620-00 (a BH candidate). CXOU J132518.2-430304 is an example of a LMXB also with a suspected stellar-sized BH.⁴⁴⁰ They produce X-ray bursts, not regular pulses. The stars being low mass must be in close proximity, so that orbital periods where known tend to be 11.4 min to 33.5 d (IMA2 p. 697). These are rare—there were 150 known LMXBs in a 2001 article⁴⁴¹ and 100 are listed in a NASA May 2012 catalog.⁴⁴²
- **Massive X-ray Binaries MXRBs:** The companion star is higher mass, often a giant O or B star (that may have survived a supernova), though often only a lower limit is known. Typically, $L_{\text{optical}}/L_{\text{X-ray}} > 1$. The traditionally defined MXRB mass is typically $> 10 M_\odot$.⁴⁴³ About 1/2 are X-ray pulsars. The X-rays are

⁴³⁵ X-ray bursters XRBs:

- Péter Mészáros, *The High Energy Universe: Ultra-High Energy Events in Astrophysics and Cosmology*, Cambridge University Press, 2010, p. 94–95
- http://en.wikipedia.org/wiki/X-ray_burster

⁴³⁶ ibid.

⁴³⁷ <http://www.sron.nl/~jeanz/bursterlist.html>

⁴³⁸ http://imagine.gsfc.nasa.gov/docs/ask_astro/answers/961213b.html

⁴³⁹ http://en.wikipedia.org/wiki/Scorpius_X-1

⁴⁴⁰ Sivakoff GR, et al, “A Transient Black Hole Low-Mass X-Ray Binary Candidate in Centaurus A”, *The Astrophysical Journal*, Volume 677, Issue 1, pp. L27–L30. <http://dx.doi.org/10.1086/587960>

⁴⁴¹ Liu QZ, et al, “A catalogue of low-mass X-ray binaries”, *A&A* Volume 368, Number 3, March IV 2001, <http://dx.doi.org/10.1051/0004-6361:20010075>

⁴⁴² Low-Mass X-ray Binaries (LMXB) NASA Catalog:

- <http://heasarc.gsfc.nasa.gov/W3Browse/all/ritterlmb.html> and

- <http://www.mpa-garching.mpg.de/RKcat/>

⁴⁴³ http://www-astro.physics.ox.ac.uk/~podsi/Flec_c1_4_c.pdf

harder, >15 keV. The stars are more widely spaced than in LMXBs, with orbital periods of 0.2 d to 580 d. They are mostly found in the plane of the MW, where younger stars including O and B types can be found. *IMA2* p. 698 states that they have been almost entirely associated with NSs so far, but that a stellar mass BH could also produce such X-ray bursts at the accretion disk. *IMA2* concludes that Cygnus X-1, a bright MXRB, is likely a BH candidate with a mass of about $14.8 M_{\odot}$ and a rotation rate at the event horizon of >800 Hz.⁴⁴⁴ Other articles such as *this*⁴⁴⁵ regarding CXOU J132527.6-430023 report possible BHs as the compact star in a MXRB, and these systems provide the best evidence for the existence of stellar-sized black holes. To deduce a black hole, the mass of the compact star must be $> 3 M_{\odot}$, i.e., above the Tolman–Oppenheimer–Volkoff limit.⁴⁴⁶

- **Intermediate-Mass X-ray Binaries IXRBs:** some authors also use this term, but the exact bounds seems to be fuzzy. Moreover, X-ray binaries in this range appear to be missing.

Fate of Binary Systems

As binary systems evolve, the low-mass secondary star in LMXBs typically evolves to a WD, with little perturbation of the circular orbit. The higher mass secondary in a MXRB may explode as a SN, which depending on relative masses as previously discussed, becomes a NS or BH. The result can be 2 NSs (or a WD and a NS) in an elongated orbit, or the SN can kick the primary NS out of orbit to become a solitary free NS. These free NS's, which are often pulsars, are mostly seen near the MW plane and often have high linear velocities from the kick, from 400 to $> 1000 \text{ m s}^{-1}$.

Also, to recap earlier discussion in this chapter, if a WD in a binary has a close donor companion star, the mass added to the WD may undergo fusion reactions leading to a nova, but this does not destroy the binary system. In extreme cases, the WD can exceed the Chandrasekhar limit and trigger a supernova that destroys the entire star.⁴⁴⁷

Millisecond Radio Pulsars

A binary of 2 NSs, for which 1 is a pulsar, can be detected by cyclical variation in the measured periods of the pulsar.

The first binary pulsar found, in 1974, was PSR 1913+16 (PSR B1913+16 or PSR J1915+1606). It has a spin period of only 56 ms and is discussed below.

PSR 1937+214 , discovered in 1982, has a spin period of only 1.558 ms, or a rotation frequency of 642 Hz. Despite the high rate of rotation, it is thought to be about 235 million years old. Although it is now solitary, it was probably once the primary in a LMXB, in which accretion from the secondary sped up its rotation. Its current rotation rate⁴⁴⁸ is decreasing at a rate of about $1 \times 10^{-19} \text{ s s}^{-1}$.

There appear to be 2 classes of binary pulsars. Those with high-mass companions evolve from MXRBs, whereas those with low mass companions (such as WDs) evolve from LMXBs... (The details of this section, *IMA2* p. 702 are confusing and seem inconsistent).

Black Widow Pulsars

It turns out that there are few WDs in binary systems with a NS, and there may be a reason for this. An example is PSR B1957+20, the so-called “Black Widow Pulsar” that is destroying its companion. This system

⁴⁴⁴ http://www.nasa.gov/mission_pages/chandra/multimedia/cygnusx1.html

⁴⁴⁵ Burke MJ, et al, “A Transient Sub-Eddington Black Hole X-Ray Binary Candidate in the Dust Lanes of Centaurus A”, *The Astrophysical Journal*, Volume 749, Issue 2,

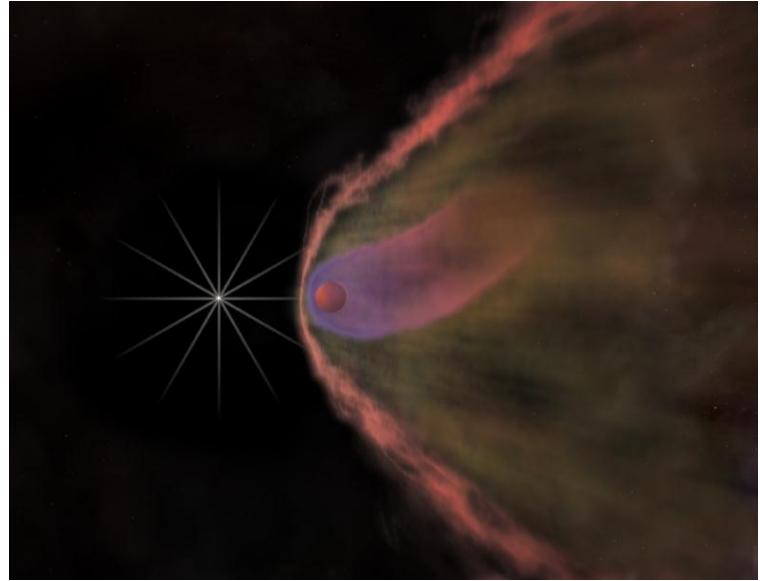
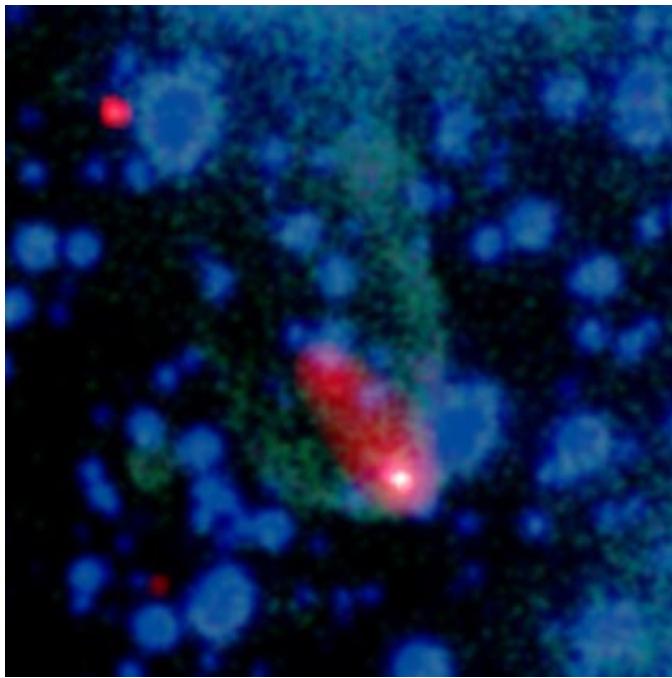
<http://dx.doi.org/10.1088/0004-637X/749/2/112>

⁴⁴⁶ http://en.wikipedia.org/wiki/Tolman%20Oppenheimer%20Volkoff_limit

⁴⁴⁷ http://en.wikipedia.org/wiki/Binary_star#Mass_transfer_and_accretion

⁴⁴⁸ http://en.wikipedia.org/wiki/PSR_B1937%2B21

is an eclipsing binary millisecond pulsar (period 1.7 ms) that orbits a brown dwarf⁴⁴⁹ or WD (*IMA2 p. 702*) companion. The orbital period is 9.2 hours and the eclipse duration is approximately 20 minutes. It is believed that the companion star is being evaporated by ablation produced by photons and charged particles from the neutron star: “In black-widow systems, a millisecond pulsar is accompanied by a low-mass companion [of a few hundredths of $1 M_{\odot}$], which is bloated and strongly irradiated by the pulsar, leading to outflows strong enough to eclipse the pulsar signal for significant fractions of the orbit. The irradiation causes strong heating on the side of the companion facing the pulsar, and, as a result, strong orbital brightness variations of the optical counterparts.”⁴⁵⁰ These authors estimate that the NS is 69 times heavier than the companion, estimated at $2.4 M_{\odot}$. “The large mass also suggests that a large amount of mass was transferred in the preceding phase as an X-ray binary, although this conclusion depends on the initial mass. However, even if that were as high as [$1.9 M_{\odot}$], our measurements suggest that the pulsar has accreted about half a solar mass.” Within a few million years, this companion will have been destroyed (*IMA2 p. 702*).



Left image: “This composite X-ray (red/white) and optical (green/blue) image reveals an elongated cloud, or cocoon, of high-energy particles flowing behind the rapidly rotating pulsar, B1957+20 (white point-like source). The pulsar, a.k.a. the “Black Widow” pulsar, is moving through the galaxy at a speed of almost a million kilometers per hour. A bow shock wave due to this motion is visible to optical telescopes, shown in this image as the greenish crescent shape. The pressure behind the bow shock creates a second shock wave that sweeps the cloud of high-energy particles back from the pulsar to form the cocoon... The Black Widow pulsar is emitting intense high-energy radiation that appears to be destroying a companion star through evaporation. It is one of a class of extremely rapid rotating neutron stars called millisecond pulsars.” The companion cannot be discerned. (NASA Chandra/ACIS = AXAF CCD Imaging Spectrometer 2001)⁴⁵¹ Distance 5000 ly, width of image 1.2 arcmin.

Right Image: An artist’s impression, viewed much closer, of the optical and X-ray emission of the NS in the “Black Widow” pulsar B1957+20 and the ablation of the companion, which orbits every 9.2 hrs. (NASA Chandra 2003)⁴⁵² “The pulsar’s orbit about the system barycentre has a radius of 0.089 light seconds projected on to the line of sight.”⁴⁵³

⁴⁴⁹ http://en.wikipedia.org/wiki/Black_Widow_Pulsar

⁴⁵⁰ van Kerkwijk MH, et al, “Evidence for a Massive Neutron Star from a Radial-Velocity Study of the Companion to the Black-Widow Pulsar PSR B1957+20”, *The Astrophysical Journal* Volume 728 Number 2, <http://dx.doi.org/10.1088/0004-637X/728/2/95>

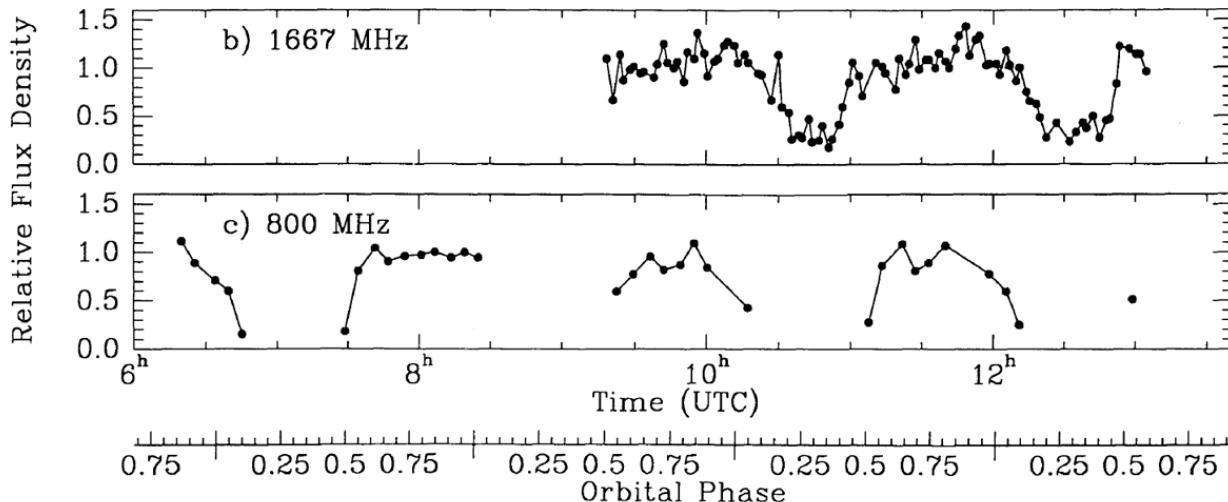
⁴⁵¹ <http://chandra.harvard.edu/photo/2003/b1957/>

⁴⁵² http://chandra.harvard.edu/photo/2003/b1957/b1957_ill2.jpg

⁴⁵³ <http://www.nature.com/nature/journal/v333/n6170/abs/333237a0.html>

Some of the ablated matter in a black widow system may condense to form one or more planets, a process suggested to explain the 3 planets orbiting the NS PSR 1257+12.⁴⁵⁴ Assuming the 6.2 ms pulsar has a standard pulsar mass of $1.4 M_{\odot}$, the derived masses of planets A, B and C are $0.03 M_{\odot}$, $4.3 M_{\odot}$, and $3.9 M_{\odot}$, respectively.⁴⁵⁵

The black widow PSR B1744-24A is another LMXB binary system in which it is thought that the secondary is being ablated by the millisecond pulsar. It is located in globular cluster Terzan 5, the NS has rotation period 11.6 ms, and orbital period of 6536 s. The pulsar signal is at least partially eclipsed for about 1/2 of each orbit, either with significant attenuation or complete absence of signal (depending on radio frequency). It is presumed that the pulsar is enveloped or at least obscured by the large cloud of ablated material arising from the companion. The eclipses are symmetrically centered at orbital phase 0.25, the stated point of inferior conjunction (Nice and Thorsett 1992). Presumably the companion is interposed between NS and the Earth at inferior conjunction. However, other sources state that binary conjunctions occur at orbital phase 0 or 0.5, so there seems to be some inconsistency in defining binary orbital phase.) The eclipsing is total at 800 MHz but only partial at 1667 MHz. This pulsar has a small but negative spin period derivative = $-1.9 \times 10^{-20} \text{ s s}^{-1}$, meaning that it “spinning up” slightly (period is decreasing).⁴⁵⁶ The findings are consistent with



Graphs of flux density of radio signals from PSR B1744-24A at 1667 MHz and 800 MHz showing partial eclipsing at 1667 MHz (2 orbits shown) but total eclipsing at 800 MHz (4 orbits).⁴⁵⁷

Double Neutron Star Binaries and Tests of General Relativity

Many double neutron star binaries have a pulsar as one of its NSs. The other NS in double NS binaries is usually not also a pulsar, and is detected only by its orbital effects.

Currently the only known exception is the **double pulsar** PSR J0737-3039 A/B,⁴⁵⁸ which has two NS pulsars of spin periods 23 ms and 2.8 s and masses of $1.4 M_{\odot}$ and $1.25 M_{\odot}$. The weaker pulsar is detectable when the stronger is eclipsed.⁴⁵⁹ (see also IMA2 p. 705, but the spin periods are a little confusing.)

The first Double Neutron Star Binary discovered was the **Hulse-Taylor pulsar, PSR 1913+16** (aka PSR B1913+16, PSR J1915+1606). It was discovered in 19744 by Russell Alan Hulse and Joseph Hooton Taylor,

⁴⁵⁴ http://en.wikipedia.org/wiki/PSR_B1257%2B12

⁴⁵⁵ Konacki M and Wolszczan A, “Masses and Orbital Inclinations of Planets in the PSR B1257+12 System”, *The Astrophysical Journal*, 591:L147-L150, 2003 July 10, <http://dx.doi.org/10.1088/0004-637X/728/2/95>

⁴⁵⁶ Nice DJ and Thorsett SE, “Pulsar PSR 1744-24A - Timing, eclipses, and the evolution of neutron star binaries”, *Astrophysical Journal*, Part 1, vol. 397, no. 1, p. 249-259, 1992,

<http://articles.adsabs.harvard.edu/full/1992ApJ...397..249N> Graph excerpted by MCM.

⁴⁵⁷ ibid.

⁴⁵⁸ http://en.wikipedia.org/wiki/PSR_J0737-3039

⁴⁵⁹ <http://www.rigel.org.uk/work/pulsar.html>

Jr.. Their analysis of the system, which strongly suggested that it was losing energy by giving off gravitational waves in accordance with the general theory of relativity, earned them the 1993 Nobel Prize in Physics.⁴⁶⁰ The rotational period of the pulsar NS is only 56 ms, and the orbital period is 7.75 hr. The 2 NSs have a mass of about $1.4 M_{\odot}$ each.

The rate of decrease of semimajor axis (orbital shrinkage) is 3.5 m y^{-1} , and the rate of *decrease* of orbital period is $0.0000765 \text{ sec per year or } -2.4 \times 10^{-12} \text{ s s}^{-1}$. Therefore, the orbital angular velocity of this binary system is gradually speeding up as a result of shrinkage (inspiralling) of the orbit, precisely as predicted by General Relativity with emission of gravitational quadrupole radiation. These results are therefore considered to be a strong indirect confirmation of the existence of gravitational radiation. (*IMA2* p. 704-5 and *here*⁴⁶¹) Here, the total power of gravitational radiation is estimated at $7.35 \times 10^{24} \text{ W}$, or about $0.02 L_{\odot}$.⁴⁶² The total orbital energy of the bound system correspondingly decreases (becomes more negative), even though the orbital period decreases and the NSs orbit about each other more rapidly in angular velocity. The NSs will eventually merge in about 300 million years.⁴⁶³ The periastron separation is only 1.1 solar radii, and the rate of periastron angular shift is $4.23^\circ \text{ yr}^{-1}$, much faster than Mercury's precession⁴⁶⁴ of the perihelion. This periastron angular shift rate is also predicted by General Relativity. (*IMA2* p. 703-4)

Other double NS binary star systems include PSR 2127+11A M15 C in globular cluster M15),⁴⁶⁵ B1534+12, and J1756–2251.⁴⁶⁶

Double neutron star pulsar binaries are superb natural laboratories in which to test the predictions of General Relativity to a very high precision.⁴⁶⁷ I have not included the formulas involved, but many are included on *IMA2* p. 704-5 or the Wikipedia page on gravitational radiation. I hope someday to be able to study the fascinating subject of General Relativity with more than superficial understanding.

Short-Hard Gamma Ray Bursts (GRB)

These are mentioned also in Chapter 15. **Short-Hard GRBs** are the result of compact object and are $< 2 \text{ s}$ in duration, averaging 0.3 s. In other words, they are associated with **NS-NS mergers** or **NS-black hole mergers** but not with core-collapse SNe (the latter are associated with long-soft GRBs). In either case, a single BH results (NASA 2005).

Perhaps the Hulse-Taylor double neutron star binary system will someday generate a Short-Hard GRB (*IMA2* p. 706).

⁴⁶⁰ http://en.wikipedia.org/wiki/PSR_B1913%2B16

⁴⁶¹ http://en.wikipedia.org/wiki/Gravitational_wave

⁴⁶² http://en.wikipedia.org/wiki/PSR_B1913%2B16

⁴⁶³ Hulse-Taylor binary pulsar PSR 1913+16:

- <http://www.astro.cornell.edu/academics/courses/astro201/psr1913.htm>
- <http://www.johnstonsarchive.net/relativity/binpulsar.html>

⁴⁶⁴ Mercury Perihelion Precession:

The total observed precession of the Perihelion of Mercury is 574.6 arc-seconds per century relative to the International Celestial Reference Frame ICRF, of which 43 arcsec is attributable to the effects of General Relativity. The ICRF (see <http://rorf.usno.navy.mil/ICRF/>) is a quasi-inertial reference frame centered on the barycenter of the solar system and defined by the measured radio positions of 212 extremely distant extragalactic objects, mostly quasars, yielding a positional accuracy of about 1 mas in both directions.) BepiColombo is a future mission planned to launch for Mercury in 2015. It is planned to test the general relativity theory by measuring the parameters gamma and beta of the parameterized post-Newtonian formalism with high accuracy.

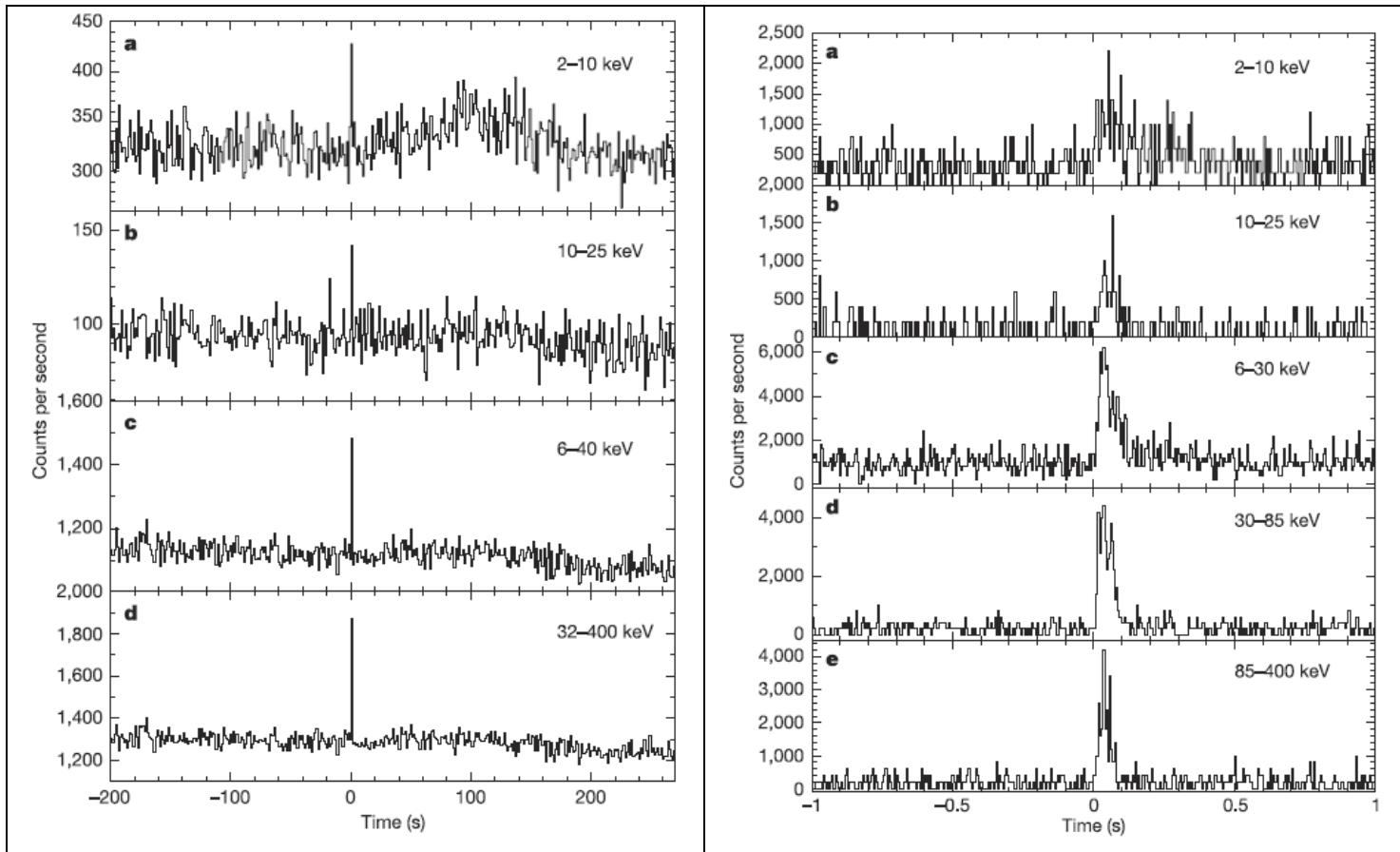
- <http://sci.esa.int/science-e/www/area/index.cfm?fareaid=30> (ESA description of BepiColombo)
- <http://en.wikipedia.org/wiki/BepiColombo>
- http://en.wikipedia.org/wiki/Parameterized_post-Newtonian_formalism
- <http://prd.aps.org/abstract/PRD/v66/i8/e082001> (Testing general relativity with the BepiColombo radio science experiment)

⁴⁶⁵ <http://adsabs.harvard.edu/full/1996IAUS..165..279D>

⁴⁶⁶ <http://arxiv.org/abs/astro-ph/0508626>

⁴⁶⁷ http://en.wikipedia.org/wiki/Tests_of_general_relativity

GRB 050709: A Short-Hard GRB was observed on July 9, 2005 (GRB 050709) which is helping to clarify the nature of these transient events.⁴⁶⁸ NASA's High Energy Transient Explorer (HETE-2) satellite was used to make the initial discovery of the GRB. Follow-up observations of the HETE-discovered event were made using ground-based telescopes, as well as the Chandra X-ray Observatory and Hubble Space Telescope. "A team led by MIT's George R. Ricker discovered a short GRB, designated GRB050709, lasting only 70 milliseconds on July 9th [2009]. The GRB was detected by NASA's High-Energy Transient Explorer (HETE-2), the first satellite dedicated to GRBs. "This particular short burst provides a long-sought nexus, enabling detection of the prompt emission and its afterglow, from the gamma-ray band to the optical, for the very first time... HETE's accurate localization of the burst allowed NASA's Chandra X-ray Observatory, Hubble Space Telescope and ground-based telescopes to identify the burst's X-ray afterglow, and, for the first time, its optical afterglow, which provided the clues needed to track the burst to its host galaxy. The distinctive signature is that of two neutron stars or a neutron star and a black hole merging, followed by a colossal explosion [rather than a supernova]."⁴⁶⁹



Left Graphs: The short X-ray and gamma-ray spike associated with GRB 050709 (July 9, 2005), shown in multiple X-ray and gamma ray energy ranges over about 470 s. An X-ray afterglow bump is also apparent in 2-10 keV. "Time history observed by WXM [HETE-2 Wide Field X-ray Monitor] in the 2-10 keV energy band (a) and in the 2-25 keV energy band (b); time history observed by FREGATE [HETE-2 French Gamma Telescope] in the 6-40 keV energy band (c) and in the 30-400 keV energy band (d). The event is a short-hard [X-ray] spike of duration $T_{90} = 220 \pm 50$ ms in the 2-25 keV energy band and 70 ± 10 ms in the 30-400 keV energy band, followed, ~25 seconds later by a long-soft bump of duration $T_{90} = 130 \pm 7$ s in the 2-25 keV energy band..."

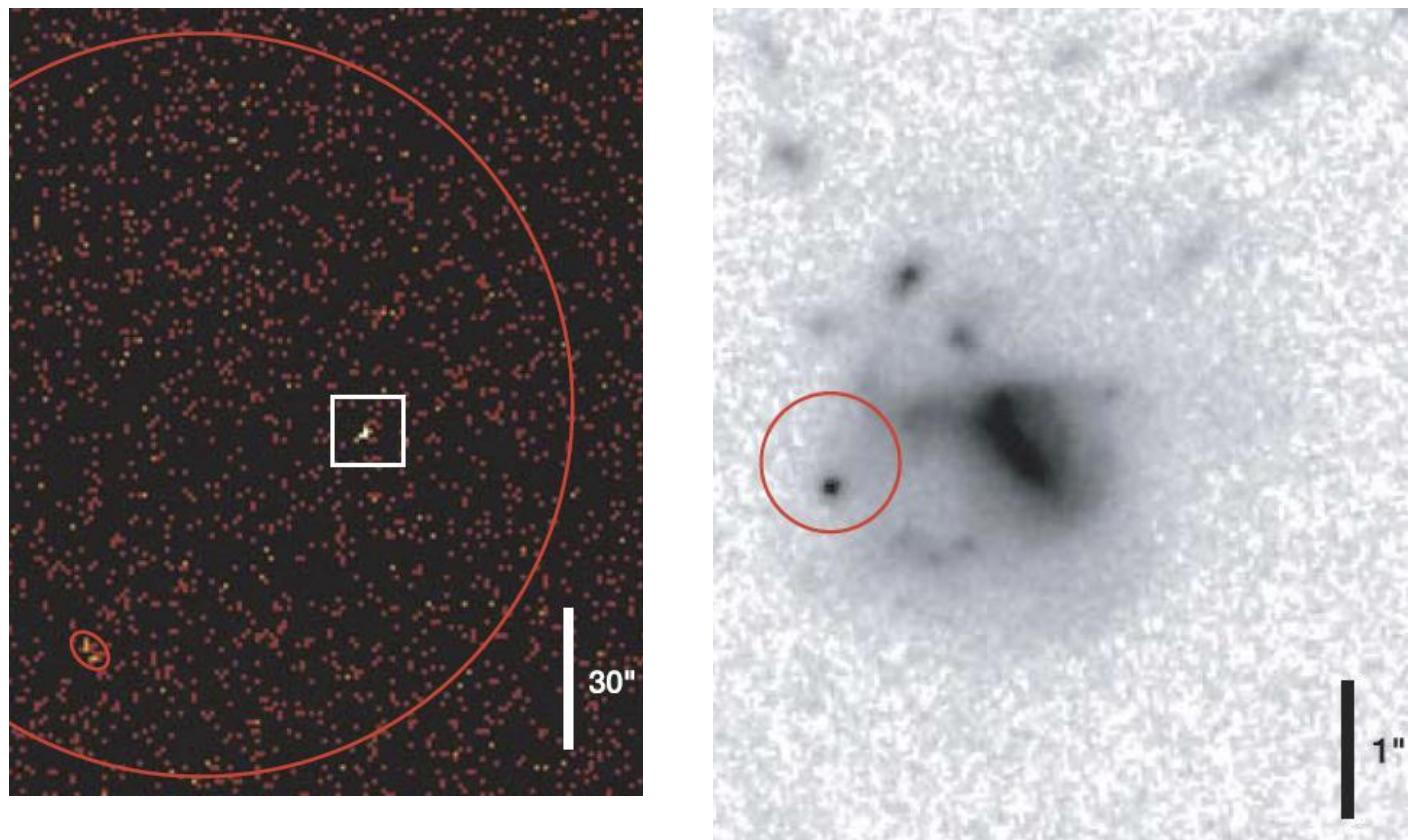
Right Graphs: Same, shorter time scale. "Time histories observed by the WXM in the 2-10 keV

⁴⁶⁸ http://www.nasa.gov/mission_pages/swift/bursts/short_burst_oct5.html

⁴⁶⁹ <http://web.mit.edu/newsoffice/2005/gamma-ray.html>

energy band (a) and the 10–25 keV band (b); and by FREGATE in the 6–30 keV energy band (c), the 30–85 keV energy band (d), and the 85–400 keV energy band (e), plotted in 5 ms time bins. The pulse has a duration of $T_{90}=220\pm50$ ms in the 2–25 keV energy band and 70 ± 10 ms in the 30–400 keV energy band, and exhibits no detectable emission before $T = 0$ or after $T=400$ ms, confirming the short, hard nature of the pulse.” (Villasenor 2005)⁴⁷⁰

The location of the X-ray and optical afterglow was initially found to lie in a circle of 90% probability having radius of 1.34 arcminutes established by the Soft X-Ray Camera (SXC) for the GRB, but the localization was subsequently improved, as noted below.



Left X-ray Image (detail): “HST and Chandra X-ray Observatory images of the afterglow and environs of GRB 050709. ... The Chandra (0.3–8.0 keV) image of the field from our observation of 2005 12.5 July UT [3 days after the July 9 GRB]. The large red circle is the HETE localization region, 81 arcsec in radius... A red ellipse indicates [an unrelated object]... The bright point source in the boxed region [box is about 1 arcmin on a side] is the afterglow of GRB 050709... Our Chandra [X-ray] afterglow candidate was found to be coincident with a point-like optical source....” Apparently the bright white spot is **X-ray afterglow**, not optical.

Right Image: “Close-up of the region surrounding the X-ray afterglow, in a co-addition of all our HST [thus optical] data; the red circle is the Chandra localization region, 0.5 arcsec in radius. A point source is visible within this region; the source is observed to fade over the course of our HST observations [see images below], and we identify it as the **optical afterglow** of GRB 050709. The irregular galaxy to its west [i.e., to the right] is the proposed $z = 0.16$ host galaxy... ”

⁴⁷⁰ Villasenor JS et al, “Discovery of the short g-ray burst GRB 050709”, *Nature* Vol 437, 6 October 2005, doi:10.1038/nature04213

We find [in] a best-fit solution ... the afterglow's 1.38-arcsec offset from the brightest region of this galaxy corresponds to 3.8 kpc in projection.”⁴⁷¹

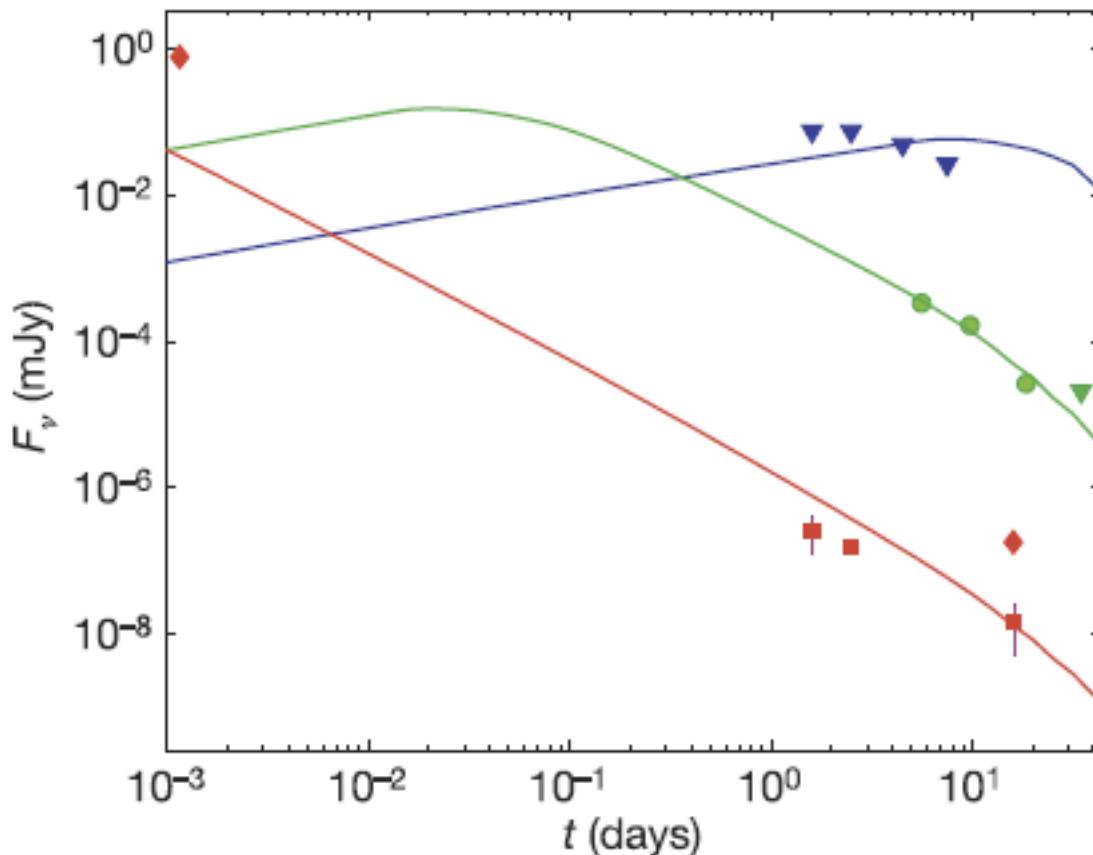
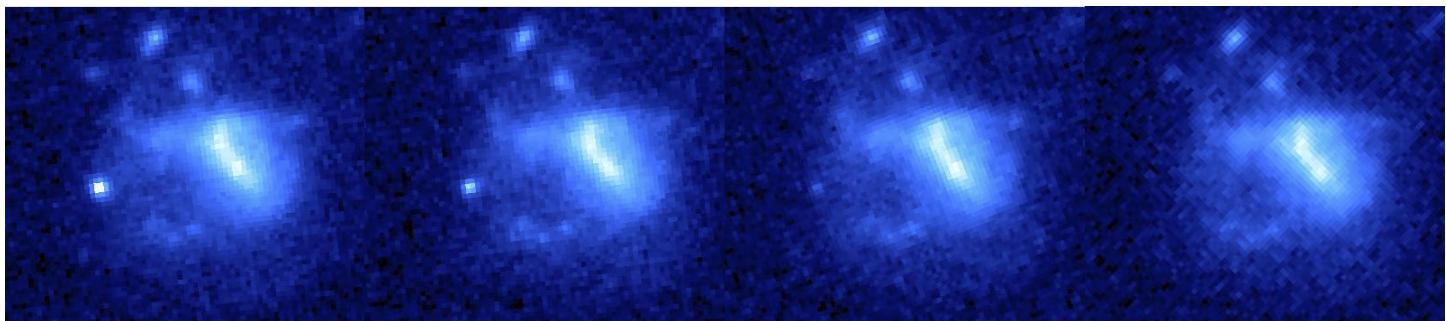


Image above: Observations of the GRB 050709 afterglow and illustrative models (curves).

X-ray = red, Optical = green, Radio = blue.

“...These multi-wavelength observations can be marginally accommodated within a standard external-shock afterglow model...” (see Fox 2005 for details) Note the log-log scaling.



Optical afterglow imaging: “These Hubble Space Telescope images show the fading afterglow and host galaxy of the HETE short burst of July 9, 2005. The images are taken 5.6, 9.8, 18.6, and 34.7 days after the burst, respectively. The bright, point-like afterglow is located to the left, and fades away over the course of the month following the burst. The colors indicate the intensity of red light (814 nm) as seen by the Advanced Camera for Surveys instrument on HST.”⁴⁷² The host galaxy is described as a bright elliptical galaxy.⁴⁷³

⁴⁷¹ Fox DB, et al, “The afterglow of GRB 050709 and the nature of the short-hard g-ray bursts”, *Nature* Vol 437-6 October 2005, doi:10.1038/nature04189

⁴⁷² Images cropped and assembled by MCM, from http://www.nasa.gov/mission_pages/swift/bursts/short_burst_oct5.html

⁴⁷³ ibid.

The Solar System & Planetary Systems (Chapters 19 through 23, omitted)

These chapters have not yet been read in my astronomy courses and are not summarized here.

The Milky Way Galaxy (Chapter 24)

History of Discoveries

The **Milky Way Galaxy (MW)** is named after its milky appearance. In fact, the word *galaxy* derives from Hellenistic Greek γαλαξίας (short for γαλαξίας κύκλος=milky circle) which came from ancient Greek ... γάλα=milk (OED). It is our galaxy, and therefore often called simply *the galaxy*. Galileo in 1610 first observed that it is made up of myriad stars. Immanuel Kant (1724–1804) in 1755 suggested it is a large stellar disk, with our Sun only a component of it, and by extrapolation that other “elliptical nebulae” might be additional extremely distant *island universes*.

William Herschel (1738–1822) produced the first attempts to map the MW 1785. The model of Jacobus Kapteyn (1851–1922) assumed the galaxy is a flattened spheroidal system and estimated stellar number density. Harlow Shapley (1885–1972) came up with a much larger (in fact excessively large) estimate of galaxy size, about 100 pc diameter, based on analysis of globular clusters, and he placed the Sun farther from the center. Early estimates of sizes were in error due to presence of unrecognized or hard to quantify *interstellar extinction*. Its effects were apparent for instance from the paucity of globular clusters within $\pm 10^\circ$ of the galactic plane.

Interstellar Extinction

This topic has been previously introduced. In the Milky Way disk, the typical extinction rate of visible light is about 1 magnitude kpc⁻¹, though obviously worse in dust lanes. The degree of extinction can be estimated by the amount of interstellar reddening. Extinction is much less for infrared.

Differential and Integrated Star Counts

Modern astronomy places great emphasis on large datasets of stars (such as SDSS,⁴⁷⁴ which by 2012 is said to have made observations on 500 million objects). The datasets are categorized by spectral types, size, luminosity, etc. The data are superimposed on models of star distributions, extinction, variations in composition, etc. that are iteratively fine-tuned until a best fit is found.

A **Differential Star Count A_m** expresses the number of stars visible in selected regions of the sky counter over a specified apparent magnitude scale. More precisely, IMA2 p 880 gives a defining formula for A_m , the number of stars having attribute(s) S and with an absolute magnitude between M and M + dM that are found within a particular solid angle Ω and that have apparent magnitude between m and m + dm.

An **Integrated Star Count N_M** expresses the number of stars having attribute(s) S and with an absolute magnitude between M and M + dM that are found within a particular cone defined by a solid angle Ω and extending from the observer to a distance d. The distance range criterion is commonly replaced by an apparent magnitude range to give an alternate form of the Integrated Star Count, \bar{N}_M , where the bar designates a function of m rather than of d.

⁴⁷⁴ http://en.wikipedia.org/wiki/Sloan_Digital_Sky_Survey

Recent and Future Astrometric Surveys of the Milky Way and Other Galaxies

According to ZXI lecture 2 and an article he co-authored⁴⁷⁵, use of the all-important SDSS dataset in a photometric survey employing **color-color diagrams** (for example, $g-r$ plotted against $u-g$), yields valuable automated data conveying much information. In the following image from that article, a locus of binary stars not previously identified was demonstrated:

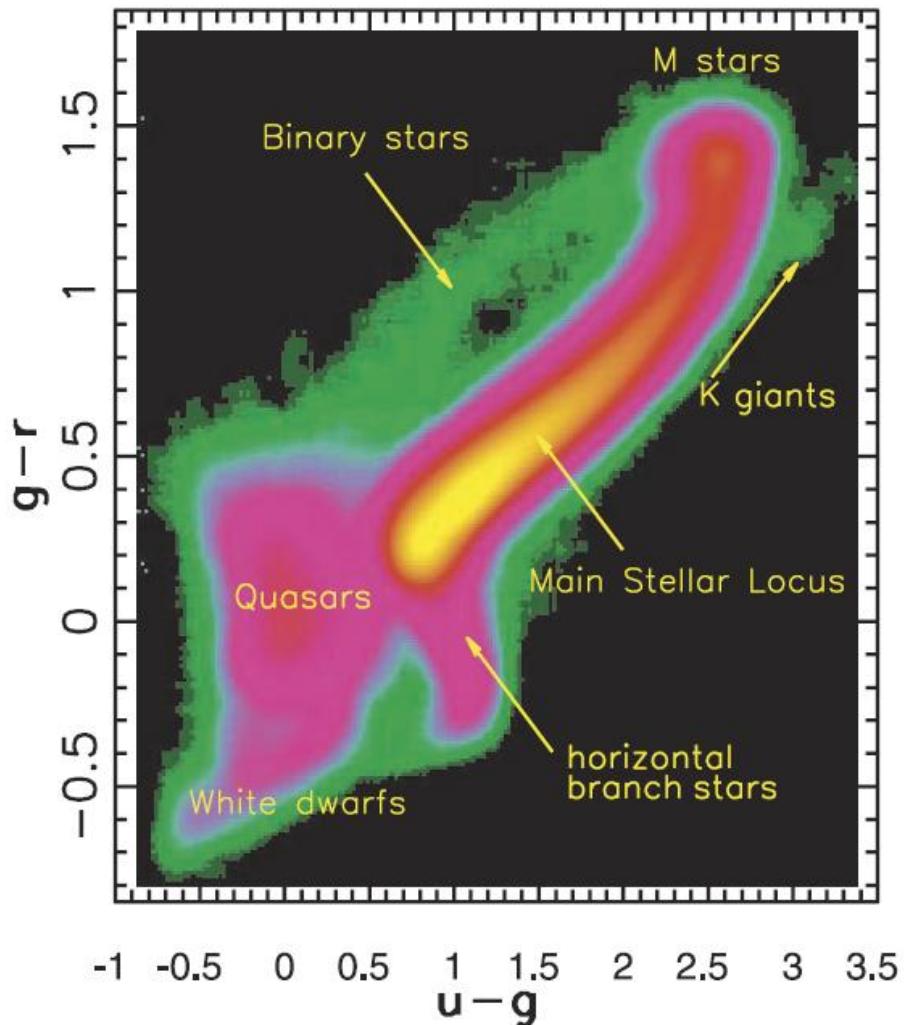
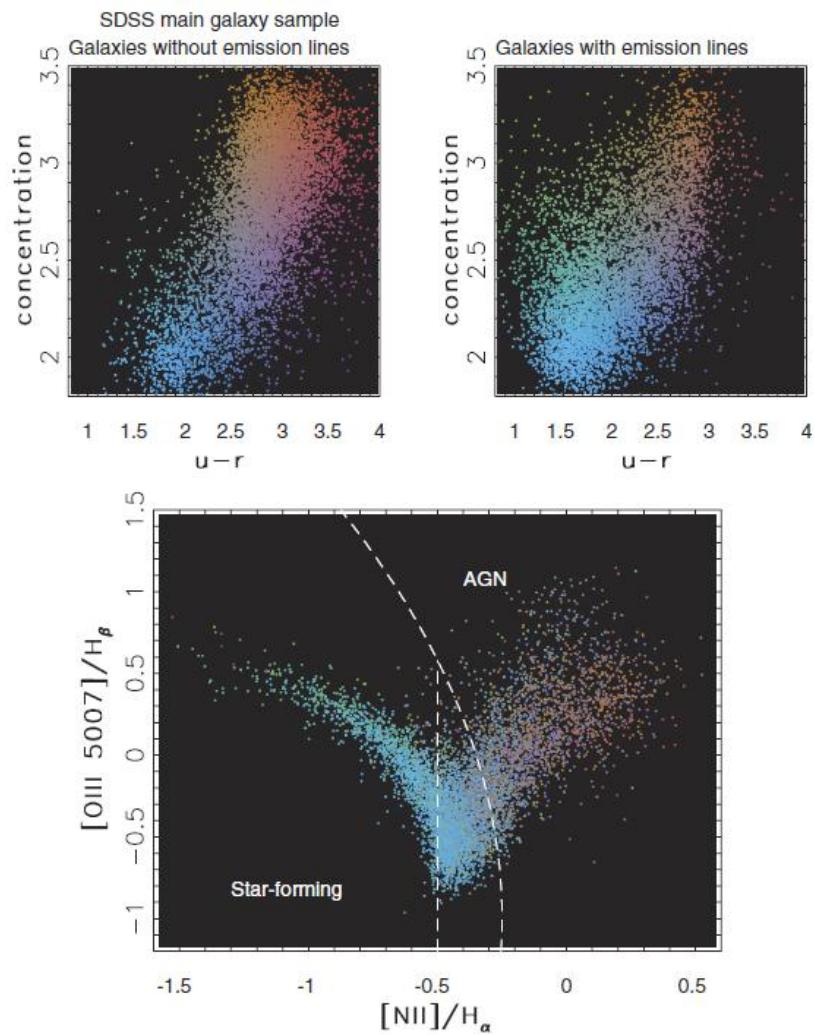


FIG. 1.—Number density, displayed on a logarithmic scale, of ~ 1.99 million stars with $u < 20.5$ from SDSS Data Release 1 in the $g-r$ vs. $u-g$ color-color diagram (increasing from green to red to yellow). The most prominent features are the main stellar locus and the clump of low-redshift ($z < 2.3$) quasars, as marked. Other notable features include the locus of white dwarfs, horizontal branch stars (also including blue stragglers and RR Lyrae stars), and solar-metallicity K giants. The fainter feature colored green (*above and to the left of the main locus*) is the locus of $\sim 1,000$ binary stars. The properties of this locus are consistent with a distribution of M dwarf–white dwarf pairs with varying luminosity ratios. The root-mean scatter of stars about this locus is only ~ 0.1 mag.

⁴⁷⁵ Smolčić V, Ivezić Z [ZXI], et al, “A Second Stellar Color Locus: a Bridge from White Dwarfs to M stars”, The Astrophysical Journal Letters Volume 615 Number 2, 2004.

In a 2008 review article ZXI also co-authored, the SDSS dataset of the Milky Way was used to establish photometric distances to map 3-D number densities of stars in tomography-like cross sections. The authors show the general structure of the MW, with a thin disk transitioning to thick disk, an oblate halo, and a number of localized overdensities including a “Virgo overdensity” in the halo toward the constellation of Virgo. (The locations studied only covered “regions more than 3 – 4 kpc from the Galactic center, where the bulge contribution is negligible.”)⁴⁷⁶

The following images also help to demonstrate the value and future potential of computer-intensive number crunching of dataset like the SDSS for galaxies. Although I have not explored or understood the details, ZXI states (Lecture 3) that in the upper plots (of “concentration index” vs. $u-r$), the galaxies without emission lines (left image) are concentrated in the upper right red region and represent ellipticals, whereas the galaxies with emission lines (right image) are concentrated in the lower left blue region and represent spirals and AGNs. In the lower image, a BPT diagram (Baldwin, Phillips & Terlevich 1981)⁴⁷⁷ plots luminosity ratios including two different forbidden transitions ($[O\ III]$ and $[N\ II]$) for emission line galaxies. The article, which ZXI co-authored, states “Emission-line galaxies can be separated into three groups according to their position in the BPT diagram: AGNs, star-forming and ‘unknown’, using the separation boundaries outlined by the dashed lines.”⁴⁷⁸ This technique clearly has value to help in identifying AGNs in an automated manner.



⁴⁷⁶ Jurić M, Ivezić Ž [ZXI], et al, “The Milky Way Tomography with SDSS. I. Stellar Number Density Distribution”, *The Astrophysical Journal* Volume 673 Number 2, 2008

⁴⁷⁷ Baldwin JA; Phillips MM; Terlevich R, “Classification parameters for the emission-line spectra of extragalactic objects”, *Astronomical Society of the Pacific, Publications*, vol. 93, Feb.-Mar. 1981, p. 5-19.
⁴⁷⁸ Obrić M, Ivezić Ž [ZXI], et al, “Panchromatic properties of 99 000 galaxies detected by SDSS, and (some by) ROSAT, GALEX, 2MASS, IRAS, GB6, FIRST, NVSS and WENSS surveys”, *Mon. Not. R. Astron. Soc.* 370, 1677–1698 (2006)

Professor Ivezić is designated System Scientist for the upcoming **Large Synoptic Survey Telescope (LSST)** project,⁴⁷⁹ which should provide further riches from surveying several billions stars. It will be located at Cerro Pachón in Chile and is scheduled to see first light in 2019. Its website states, it “will produce a 6-band (0.3-1.1 micron) wide-field deep astronomical survey of over 20,000 square degrees of the southern sky using an 8.4-meter ground-based telescope. Each patch of sky will be visited about 1000 times in ten years. The LSST leverages innovative technology in all subsystems: the camera (3200 Megapixels)..., telescope ..., and data management (30 terabytes of data nightly, nearly instant alerts issued for objects that change in position or brightness...)”

The space-based **GAIA Mission** of the ESA is also highly ambitious and eagerly anticipated. It will be placed in orbit at the L₂ Lagrangian point, which “lies at a distance of 1.5 million kilometres from the Earth in the anti-Sun direction and co-rotates with the Earth in its 1-year orbit around the Sun.”⁴⁸⁰

Perhaps I will have more time eventually to delve deeper into these new and impressive areas of research.

Milky Way Fundamentals, Morphology, and Coordinates

Our galaxy is hard to study because we are embedded in it, there is substantial extinction close to the galactic plane, and because we are on a moving platform moving in a complex way relative to the galactic center.

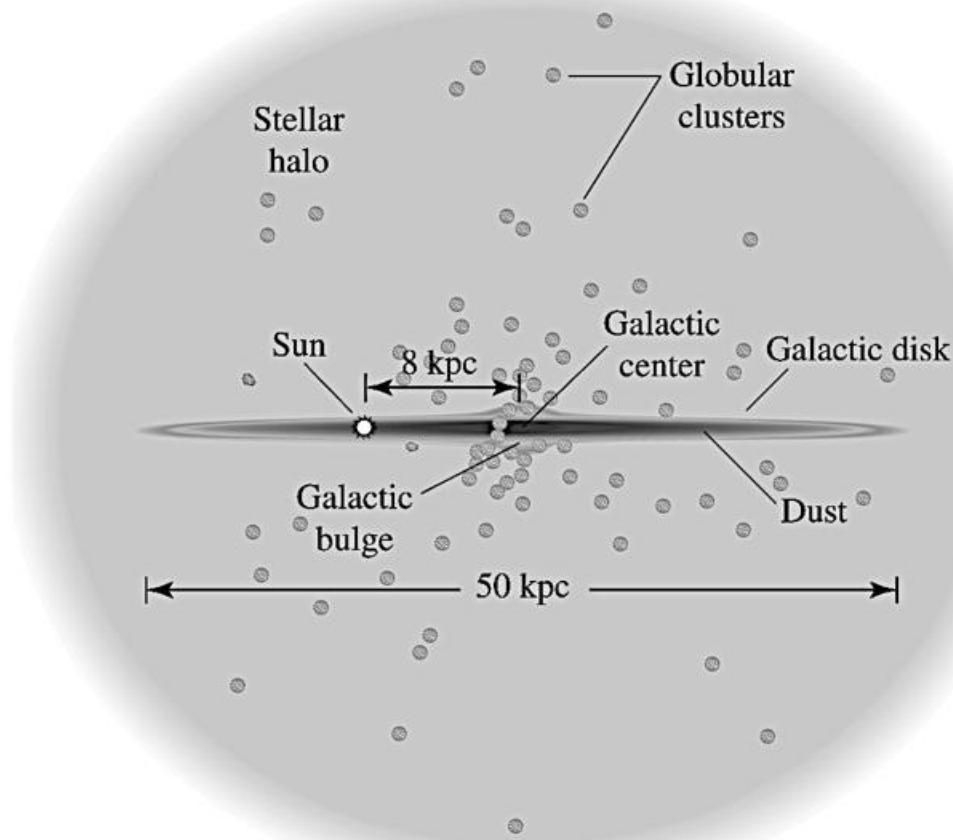
The Milky Way is a **barred spiral galaxy**, type **SBc** (or **SBbcI-II**). It has approximately 300×10^9 stars, a total mass of $1.0 \text{ to } 1.5 \times 10^{12} M_{\odot}$, and a bolometric luminosity of $2 \text{ to } 3 \times 10^{10} L_{\odot}$ (based on an estimate of total luminosity $L_{\text{MW}} = 10^{44}$ ergs/s). The oldest known star is 13.2 Gy.⁴⁸¹ The average MW stellar mass is about $0.7 M_{\odot}$, thus dwarf stars smaller than the Sun predominate (*IMA2* p. 887).

As shown in the diagram to follow, the full galactic disk (stars, dust, and gas) is about 50 kpc = 163,000 ly in diameter, although the disk is thought to be somewhat elliptical (0.9 minor/major) with the Sun near a major axis.

⁴⁷⁹ <http://www.lsst.org/lsst/>

⁴⁸⁰ <http://sci.esa.int/science-e/www/object/index.cfm?fobjectid=28820>

⁴⁸¹ Anna Frebel et al, “Discovery of HE 1523-0901, a Strongly r-Process-enhanced Metal-poor Star with Detected Uranium”, *The Astrophysical Journal Letters* Volume 660 Number 2, 2007 ApJ 660



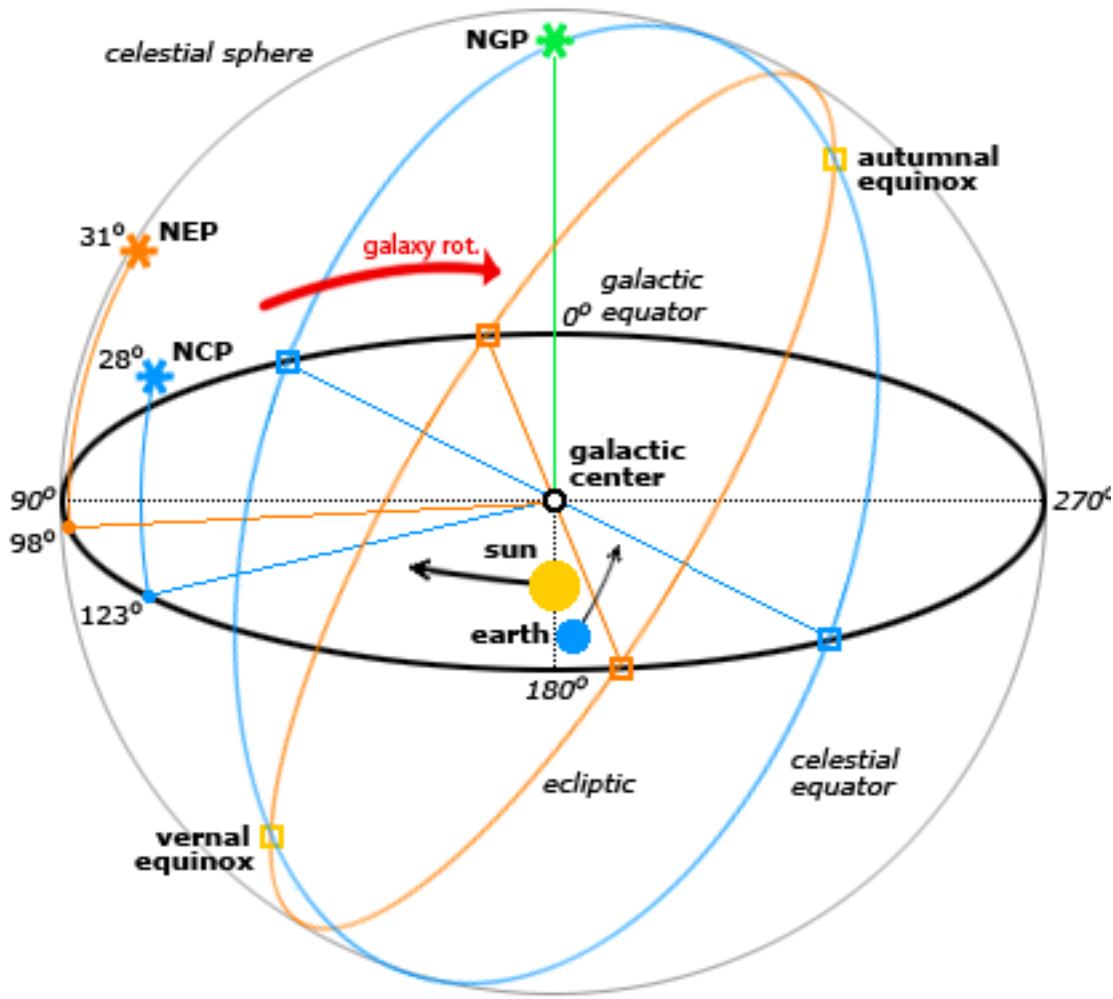
Gross Morphology of the Milky Way⁴⁸²⁾

Galactic Coordinate System and Cylindrical Coordinate System

The central plane of the Milky Way is oriented at 62.87° with respect to the plane of the celestial equator. In a particular view, the **north celestial pole NCP**, **north ecliptic pole NEP**, and **north galactic pole NGP** are all in the same region of the celestial sphere. Looking down from the NGP, the Milky Way is seen to rotate clockwise with respect to the celestial sphere (the opposite of the Earth's rotation about the Sun).

Galactic coordinate angles are measured with the Sun at the vertex of the angle (even though the Sun is not at the galactic center). The axis pointing from the Sun to the center of the galaxy, ~8 kpc away, defines 0° galactic longitude ($\ell=0$). The galactic longitude for an object is measured in an easterly direction (counterclockwise looking down from NGP) from the direction of the 0° galactic longitude to the intersection with the galactic equator of a plane oriented perpendicular to the galactic plane and passing through the Sun and the object (but *not* usually through the NGP!). The galactic latitude b for an object is the angle in degrees, positive for north and negative for south, between the object and the galactic plane with Sun at the vertex and lying again in a plane holding the Sun and the object and perpendicular to the galactic plane.

⁴⁸² from <http://faculty.physics.tamu.edu/papovich/courses/milkyway2.pdf> , which derives from image in IMA2 p. 883



Comparison of Galactic, Ecliptic, and Celestial Coordinate Systems:
 NCP=north celestial pole (blue), NEP=north ecliptic pole (gold), NGP=north galactic pole (green)
 Celestial equator (blue), Ecliptic (gold), galactic equator (black)
 Rotation of galaxy shown wrt fixed celestial sphere.
 (Rotation annotated by MCM)⁴⁸³

⁴⁸³ MCM modified from http://en.wikipedia.org/wiki/Celestial_coordinate_system
 Page 178 of 249

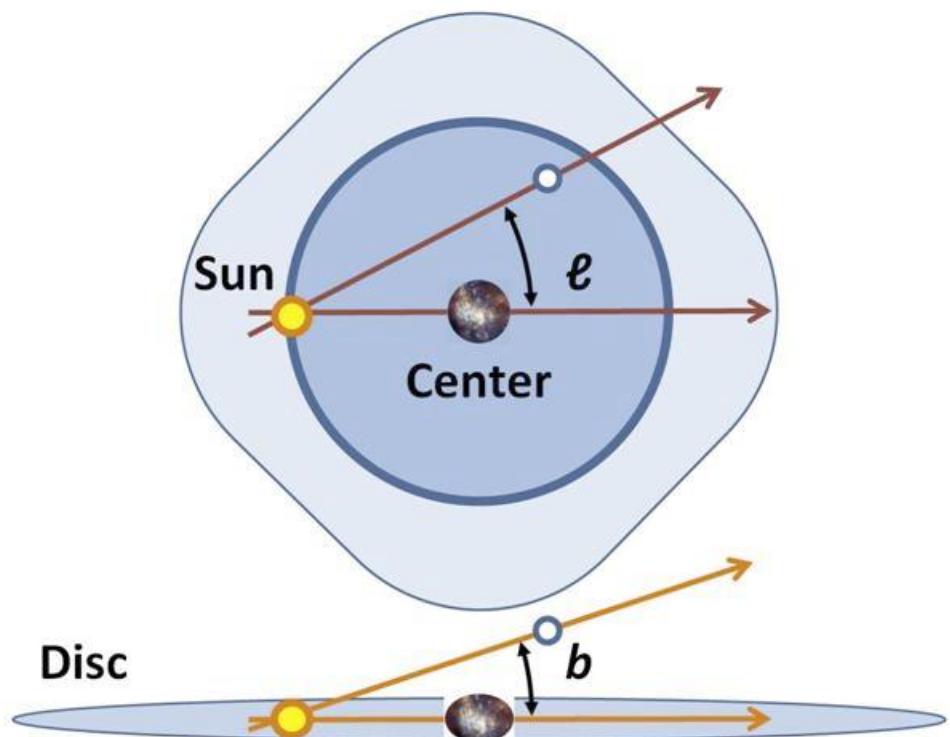


Diagram showing how latitude and longitude are measured with the Sun at the vertex.⁴⁸⁴

Galactic motions (kinematics) are measured in a cylindrical coordinate system, centered on the galactic center and using coordinates R , θ , and z . Here, θ is the azimuthal-type angle pointing in the direction of rotation and with the Sun at 0° , R is the radius from the galactic center to the object's projection on the galactic plane, and z is vertical distance from the galactic plane.

Thick and Thin Disk

The disk is subdivided into the **thin disk** and the **thick disk**, the latter being much lower in M , L , and star number density. Disk heights and lengths are expressed by **scale heights and scale lengths**. **Scale height** is defined as the vertical z distance from the central Galactic midplane over which a quantity such as the number density decreases from the typically higher central value by a factor of $1/e = 0.37$ (i.e., the final value is $1/e$ times the initial value). At the central axis of the disk, the thin disk has a scale height of $350 \text{ pc} = 1140 \text{ ly}$, whereas the thick disk is $\sim 1000 \text{ pc} = 3300 \text{ ly}$. The disk scale length in the radial direction is about 2.25 kpc . The central star number density is $n_0 \approx 0.02 \text{ stars pc}^{-3}$. *IMA2* p. 883 gives a formula for star number density $n(z, R)$ in the disk, based on these scale heights and scale length.

The thin disk consists of mostly younger stars, dust and gas. The thicker disk consists of older stars and has a stellar number density of about 9% of that in the Galactic midplane. Comparative values for M , L , R , $[\text{Fe}/\text{H}]$, Age, etc. are given on *IMA2* p. 884.

Disk stars rotate in near-circular orbits about the galactic center and stay confined to the disks.

Spiral Bands

The spiral structure so readily apparent in other galaxies (such as Andromeda M31) is more difficult to discern in the MW because we are embedded in it. The arms are the site of new star formation and exhibit more blue light as a result of greater numbers of these young, hot, bright massive stars (though not necessarily much greater number density of stars overall). Spiral arms are less conspicuous in red light. The Sun is close to but not in the Orion-Cygnus arm, aka the Orion arm or Orion spur.

⁴⁸⁴ www.astro.sunysb.edu/metchev/AST443/lecture2.pdf (Stanimir Metchev)
Page 179 of 249



"Note also that the Orion Arm is not a major spiral arm, but only an enhancement of stars and gas between the Sagittarius and Perseus arms. It should be emphasized that there are almost as many stars between the spiral arms as in the spiral arms. The reason why the arms of spiral galaxies are so prominent is that the brightest stars are found in the spiral arms. Spiral arms are the major regions of star formation in spiral galaxies and this is where most of the major nebulae are found."⁴⁸⁵

Interstellar Gas, Dust, and Peripheral Warp

The 21-cm microwave and IR emissions of CO are useful for assessing H I and H₂ in the central plane and the distant reaches of the MW. (IMA2 p. 889) At R > 12 kpc, the scale height of H I increases and a pronounced warp in the z-axis-distribution of H I becomes apparent by R = 13.6 kpc.⁴⁸⁶

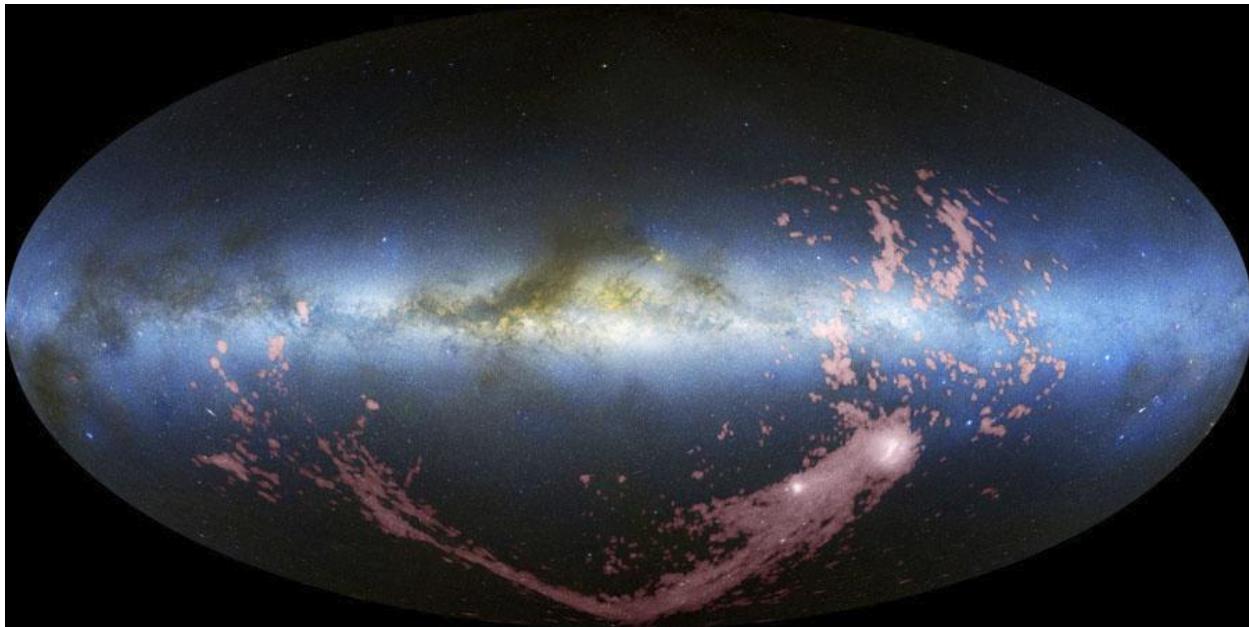
⁴⁸⁵ <http://www.atlasoftheuniverse.com/milkyway.html>

⁴⁸⁶ Burton WB, te Lintel Hekkert P, "The shape of the outer-Galaxy H I layer. I - Atlas of volume densities in cuts through a composite, galactocentric data cube", *Astronomy and Astrophysics Supplement Series* vol. 65, no. 3, Sept. 1986, p. 427-463

There is evidence for high-velocity clouds of hydrogen, and a **Galactic Fountain model** has been suggested, perhaps arising as SN ejecta. There is a very hot tenuous gas at distances up to 70 kpc from the galactic center, apparent in other galaxies from O VI absorption lines. This hot gas is estimated to have $T > 10^6$ K. By analogy to the Sun, it is called a **coronal gas**.

Disruption of Satellite Galaxies

The Magellanic Stream is a narrow but long band of H I radio emission stretching more than 180 degrees across the Southern Hemisphere and trailing the Magellanic Clouds. It is proposed that this stream is a result of tidal encounters of the Large and Small Magellanic Clouds with the MW some 200 Ma. (IMA2 p. 891)



Magellanic Stream Radio Emission Superimposed on All-sky Visible Light Image.

"Wide angle radio images—including those from the Byrd Green Bank Telescope—have shown that the Magellanic Stream is longer and older than previously thought, perhaps as old as 2.5 billion years. These observations bolster a third possible origin for the stream—that the Large and Small Magellanic Cloud galaxies [SMC and LMC] once passed so close to each other that gravitational tides triggered a burst of star formation that left the stream. Pictured above digitally superposed on a recently-completed all-sky image in visible light, the radio emission of [the] Magellanic stream is shown in false color pink extending across the sky and ending at the two Magellanic galaxies on the lower right." (NASA)⁴⁸⁷

Other possible former or current satellite galaxies that may have interacted with the MW include the Sagittarius dwarf spheroidal galaxy (leading to the Sagittarius Stream), a group of globular clusters in Canis Major, and the unusual globular cluster ω Centauri.

Galactic Bulge

This feature, which is better seen in IR, is thought to be an independent component of the MW (and of many spiral galaxies). It has a mass of about $10^{10} M_{\odot}$, and B-band luminosity of $3 \times 10^9 L_{\odot}$. and $M/L \cong 3 M_{\odot}/L_{\odot}$. The somewhat "boxy" bulge is apparent in COBE IR views (IMA2 p. 892).

As with other galactic features, the ill-defined boundaries of the bulge may be delineated by a threshold scale height (distance for quantity to drop to $1/e = 0.37$ of the highest value), and for the MW this is 100 to 500 pc depending on star type.

⁴⁸⁷ <http://apod.nasa.gov/apod/ap100125.html>

The quantity assessed to define its boundary is typically **surface brightness I**. (See earlier discussion of surface brightness) An $r^{1/4}$ relationship for I has been empirically found (*IMA2* p. 892), using the effective radius r_e and I_e as this radius as normalization values:

$$\log_{10} \left[\frac{I(r)}{I_e} \right] = -3.3307 \left[\left(\frac{r}{r_e} \right)^{\frac{1}{4}} - 1 \right],$$

here expressed in units of L_\odot per pc⁻². This is termed **de Vaucouleurs' law**⁴⁸⁸, a special case of Sersic's (Sérsic's) law with Sersic index n=4 (discussed later). By IRAS, the effective radius r_e is found to be 0.7 kpc. The chemical characteristics and metallicity of stars in the bulge are discussed at *IMA2* p. 893.

The stars have largely random orbits but with some net rotation about the Galactic center. There is a ring of gas and dust near the center. (ZXI lecture 2).

Central Bar and 3-kpc Expanding Arm

This has been found in the bulge. It has a half-length (radius) of ~4.4 kpc, is somewhat thicker in the central plane of the galaxy, and is oriented at 44° with respect to the Sun to Galaxy center line. An inner expanding arm at 3 kpc seen “near” and “far” appears to be related to the bar.⁴⁸⁹

Stellar Galactic Halo

This roughly spherical feature of visible (luminous or demonstrable) matter extends about 50 kpc from the disk, thus about the same as the diameter of the disk. It consists of globular clusters of stars and **field stars** (stars that are not part of a cluster and that often have a high velocity). Two populations of globular clusters are apparent: older metal-poor globular clusters ($[Fe/H] < -0.8$), which extend in a nearly spherical distribution, and younger clusters with $[Fe/H] > -0.8$ which occupy a flatter distribution (see *IMA2* p. 894 ff for details). Some of the distant clusters may have been captured as clusters of dwarf galaxies. It contains little gas and dust, and no recent star formation. Unlike the disk stars (which orbit around the galactic center in the disk plane), the halo stars have randomly distributed elliptical orbits in three dimensions.

Although it was previously thought that there was no obvious substructure in the halo, structural features such as the Sagittarius Stream, Virgo Overdensity, Monoceros Ring, and Orphan Stream have been identified with SDSS data (ZXI lecture 2 and *here*⁴⁹⁰).

The halo is of interest because it exhibits merger history and because the gravitational potential becomes dominated by the dark matter halo (ZXI Lecture 2b).

Total Galactic Luminosity and Mass (not including dark matter)

The estimated total MW luminosity in the B-band (blue band) is $L_{B, MW \text{ tot}}$ is $\sim 2.3 \cdot 10^{10} L_\odot$ and total bolometric luminosity, including IR, is $L_{\text{bol}, MW \text{ tot}}$ is $\sim 3.6 \cdot 10^{10} L_\odot$. (As usual, I believe that the total bolometric value of L_\odot is used for the comparison, not the blue band solar luminosity, $L_{\text{blue}, \odot}$.) The total mass M_{MW} is $\sim 9 \cdot 10^{10} M_\odot$.

Dark Matter Halo

This is inferred by the unexpectedly flat distribution of star and gas rotational velocities beyond R_0 (8 kpc) from the center, and may also contribute to the peripheral warp mentioned. It extends to 230 kpc or farther, with an empirical density distribution of

⁴⁸⁸ http://en.wikipedia.org/wiki/De_Vaucouleurs%27_law

⁴⁸⁹ http://www.cfa.harvard.edu/mmw/Far_3kpc.html

⁴⁹⁰ <http://www.sdss.org/news/releases/20060508.mergers.html>

$$\rho(r) = \frac{\rho_0}{(r/a)(1+r/a)^2}$$

where $\rho(r)$ is density at 3-D radius r , and ρ_0 and a are fitted values (*IMA2* p. 896, see later discussion). The dark matter halo comprises about 95% of the mass of the MW. The nature of this dark matter is actively debated and under investigation (and not summarized here). Possibilities include nonbaryonic weakly interacting massive particles (WIMPS), neutrinos, neutralinos, and massive compact halo objects (MACHOS, such as black holes, neutron stars, and white or brown dwarfs).

Galactic Magnetic Field

Typical MW magnetic fields (based on polarization by aligned interstellar grains and by Zeeman splitting) are 0.4 nT in the spiral arms, 1000 nT in the galactic center, and perhaps 0.04 nT in the stellar halo. By comparison, the Earth magnetic field is about 5×10^{-5} T = 50 microT = 50,000 nT = 0.5 gauss.

Despite the low field strength, the energy density from the magnetic field is comparable to that of the thermal energy, at least in the disk.

Age-Metallicity Relation

The Population I stars are relatively metal-rich with $Z \sim 0.02$, whereas Population II are metal-poor with $Z \sim 0.001$. Metallicity is also expressed by [Fe/H], and [O/H], as previously discussed, because Fe lines are readily observed in spectra. During a supernova explosion, particularly Type Ia, Fe is ejected enriching the ISM and metallicity of future stars that form. Iron content then correlates with stellar age, greater metallicity being found in younger stars (such as the Sun for which [Fe/H]=0.0 by definition), and lower metallicity is seen in old extremely metal-poor stars (with values down to -5.4). Metallicity has been used to show that thin disk stars forming ~8 Gya and more recently are younger by 2 or 3 billion years than thick disk stars, which formed ~10 - 11 Gya. (*IMA2* p. 886)

This is the age-metallicity relation, but it is not always reliable. Type IA SN do not occur until about 10^9 years after star formation, and mixing of SN ejecta in the ISM is not uniform. Core-collapse SN produce more [O/H], which may be helpful. Age estimation is further complicated by the strong dependence on distance modulus.

Kinematics of the Milky Way

(Highly abbreviated and selected)

As noted, galactic kinematic motions are measured in a cylindrical coordinate system, centered on the galactic center and using coordinates R , θ , and z . Here, θ is the azimuthal-type angle pointing in the direction of rotation and with the Sun at 0° , R is the radius from the galactic center to the object's projection on the galactic plane, and z is vertical distance from the galactic plane. (The symbol r is customarily reserved for 3-D distance to the object from the center.) Object velocity components for R , θ , and z are expressed as Π , **Θ , and Z** , respectively.

Formulas for conversion to/from the galactic coordinate system are given on *IMA2* p. 901 ff, including a graph for making these conversions.

Observations measured from the moving Earth must be corrected for peculiar motions and the Local Standards of Rest LSR. The LSR has an orbital speed of $\Theta_0 R_0 = 220$ km s⁻¹. The velocity of a star relative to the LSR is known as the star's **peculiar velocity**. The peculiar velocity is given by the $\mathbf{V} = (u, v, w)$, with components representing components V_R, V_Θ and V_z . These and other complex kinematic matters are discussed in *IMA2* p. 901-913. The Sun's own peculiar velocity relative to the LSR is the **solar motion**.

For a star, plotting one component of peculiar velocity against another, one can plot velocity ellipsoids. These can be used to show that the oldest metal-poor stars in the Galaxy have the widest range of peculiar velocities in all three coordinates, whereas younger metal-rich stars have smaller peculiar velocities and are found in

the thin disk. Stars with large $|w|$ are passing through the solar neighborhood on trajectories that will carry them to great distances above and below the disk.

The amount of variation in each component of velocity for a population of stars is expressed as the **standard deviation σ** of the velocity distribution, given by

$$\sigma \equiv \frac{1}{\sqrt{N}} \left[\sum_{i=1}^N (u - \langle u \rangle)^2 \right]^{1/2}$$

When $\langle u \rangle$ is zero, the quantity σ here computed is

$$\sigma = \sigma_u \equiv \frac{1}{\sqrt{N}} \left[\sum_{i=1}^N (u)^2 \right]^{\frac{1}{2}} = \langle u^2 \rangle^{1/2}$$

and is termed the **velocity dispersion**.

The Milky Way is moving in the direction of $\alpha = 10.5^\circ$, $\delta = -24^\circ$ (J2000, near the center of Hydra)⁴⁹¹ at 552 km/s with respect to the CMB rest frame.

The Sun is moving toward the **solar apex**, opposite from the **solar antapex**. Objects in eccentric orbits about the center are said to attain the **apogalacticon** (greatest distance from the center) and the **perigalacticon** (closest distance).

It takes about 250 million years for the Sun to make a full orbit around the Milky Way.

Differential Galaxy Rotation and Oort's Contributions

The galaxy is rotating with differential rotation, as first analyzed by Jan Oort (1900–1992). The angular velocity of an object is given by $\Omega(R) = \Theta(R)/R$ where $\Theta(R)$ is linear orbital velocity. Oort derived formulas incorporating **Oort Constants A and B** which give $v_r \cong Ad \sin 2l$ and $v_t \cong Ad \cos 2l + Bd$. (Here, l is galactic longitude, d is distance from Sun to star, see *IMA2* p. 911).

⁴⁹¹ http://en.wikipedia.org/wiki/Milky_Way

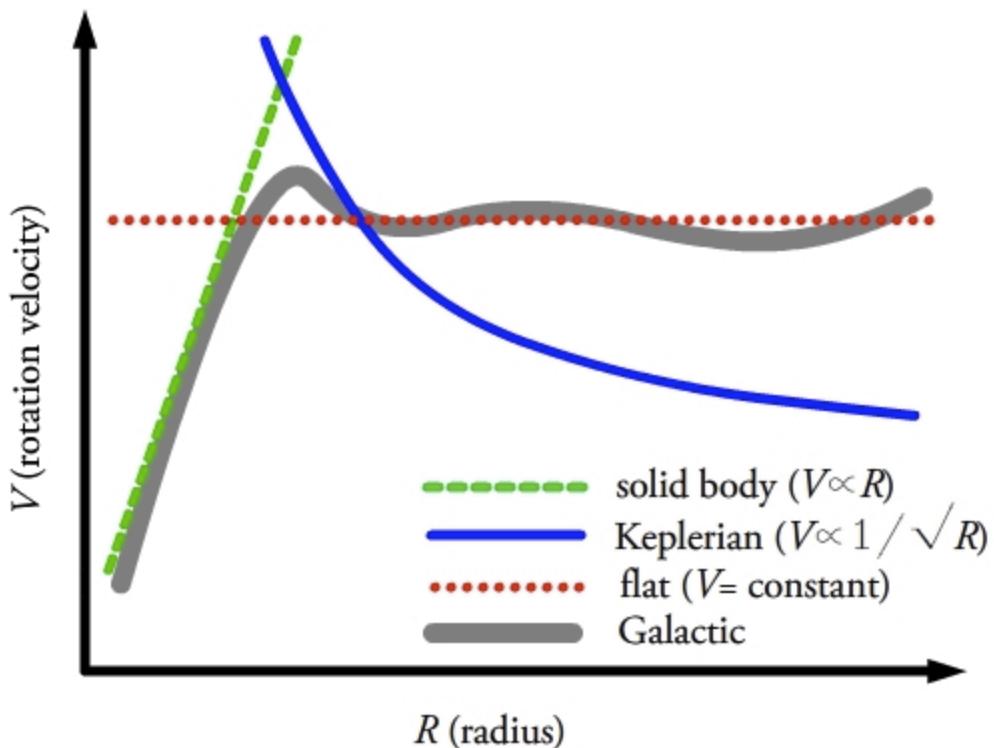


Diagram of the various rotation curves in a galaxy⁴⁹²

The Oort constants help one understand the types of possible galactic rotation. **Solid Body rotation** is self-evident, the angular velocity $\Omega(R)$ is constant with R and rotational velocity $\Theta(R)$ rises steadily with increasing R . There is no shear motion, $A = 0$, and the vorticity is just the angular rotation $B = -\Omega$. This gives the upper limit or fastest an object can rotate for a given radius. This model fits the observed data only in the inner regions of the galaxy (see graph).

In the case of **Keplerian rotation**, where $v = (GM/r)^{1/2}$, observed rotation values of Oort constants do not match the derived values, so the Keplerian model appears invalid. It can however be thought of as providing the minimum velocity an object can have in a stable orbit.

The flat rotation curve in which linear rotational velocity $\Theta(R)$ is constant and independent of radius yields Oort constants that fit remarkably close to observed data in the solar neighborhood and farther out. It provides a value intermediate between rigid and Keplerian motion.

Hydrogen 21-cm Line as a Probe of Galactic Structure

Because of light extinction, this radio or microwave line from H I is very useful in exploring the entire galactic disk. However, it is difficult to determine distance d with it. The technique is discussed on *IMA2* p. 914.

Estimating Distances in the Milky Way

The distances to MW stars for galaxy analysis are most reliably estimated with Cepheid Variables.

The **moving cluster method** is also important. Early work was done especially with the Hyades open cluster of stars, a cluster in the MW at about 46 - 47 pc.⁴⁹³ The directions of stellar motions over time may be deduced, after removing the effects of the Sun's peculiar motion. When plotted in equatorial coordinates, the star motions appear to point to a **convergent point**. The distance to the cluster may be determined from

⁴⁹² http://en.wikipedia.org/wiki/Oort_constants

⁴⁹³ http://en.wikipedia.org/wiki/Hyades_%28star_cluster%29

knowledge of the direction to the convergent point, the average radial velocity of the cluster members, and the average of their proper motions (*IMA2* p. 921).

Once a distance to a cluster is deduced, one may calibrate the main sequence. By comparing the apparent magnitudes of other cluster H-R diagram main sequences, it is possible to find the distance modulus of these clusters as well (main-sequence fitting). This technique is useful because it uses a large number of stars to reduce statistical errors (as compared to spectroscopic parallax). One may use RR Lyraes in clusters of known distance to determine intrinsic luminosities.

Secular parallax may also be used by repeating measurements over several years and taking advantage of the Sun's peculiar motion with respect to a group of stars.

The Flat Rotation Curve and Evidence of Dark Matter

Compiled data shows a relatively flat rotation speed curve for the MW beyond the Sun's galactocentric radius, here taken to be 8.5 kpc (although the error bars are wide):

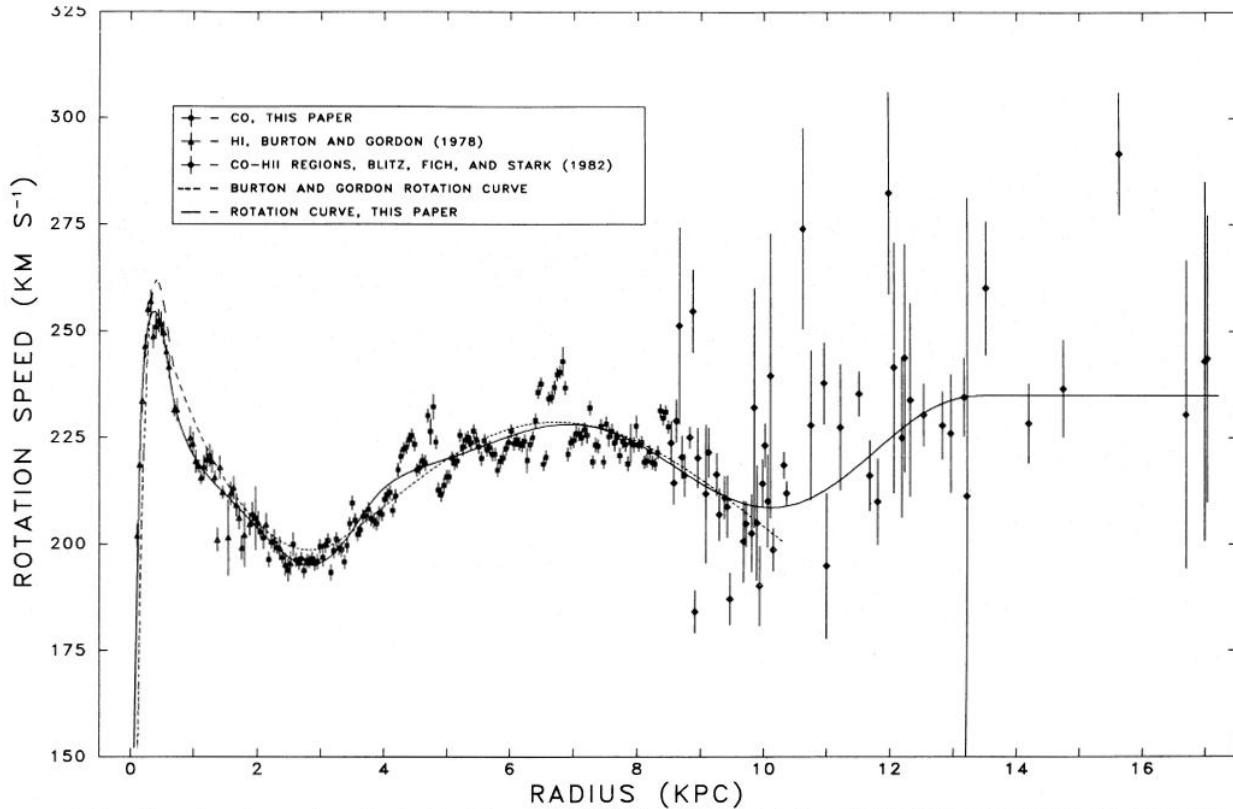


FIG. 3.—Plots of the rotation speed versus galactocentric radius. The solid lines correspond to the polynomials, and the dashed lines are the BG rotation curve. (upper panel) $(R_0, \Theta_0) = (10 \text{ kpc}, 220 \text{ km s}^{-1})$; (lower panel) $(8.5 \text{ kpc}, 220 \text{ km s}^{-1})$.

Clemens' graph (DP Clemens 1985)⁴⁹⁴ of rotation speed versus galactocentric radius, data from multiple authors plus Clemens' paper , assuming $R_0=8.5 \text{ Mpc}$, $\Theta_0(R_0)=220 \text{ km s}^{-1}$.

The near-flatness of this curve of rotation speed beyond $R \approx 8 \text{ Mpc}$ is surprising, as most of the luminosity of the Galaxy is produced by matter residing inside R_0 . Similar observations by Vera Rubin (b. 1928) and others have been made for other galaxies (such as for NGC 2998, NGC 4378, etc.) and support the existence of dark matter. (*IMA2* p. 914 – 919)

The empirical formula (also presented earlier)

$$\rho_{\text{NFW}}(r) = \frac{\rho_0}{(r/a)(1+r/a)^2}$$

⁴⁹⁴ Clemens, DP, "Massachusetts--Stony Brook Galactic Plane CO Survey: The Galactic Disk Rotation Curve", *ApJ*, 295, 422 (1985).

of Navarro, Frenk and White was proposed in 1996.⁴⁹⁵ It is applicable over a large range and tapers better at large radius (near the edge of the halo). However, it is still not unbounded to infinity, and it is possible that other galaxies may overlap and perhaps share the dark matter (*IMA2* p. 219 and my inference).

Galactic Center

The galactic center is located in the Sagittarius constellation. The IAU 1985 agreed-upon distance to the galactic center from the Sun (the **solar Galactocentric distance R_0**) is 8.5 kpc, but more recent data and the authors of *IMA2* have taken 8 kpc as the best working estimate (thus about 26,100 ly). The Sun lies about 30 pc above the central galactic plane, and is rising so that it will attain a maximum z of 85 pc above the plane in 15 million years, at which time we will have a better view of the center!

This regions is best viewed with radio, microwave, and infrared, and also X-rays and gamma rays. The K band of IR is useful, and helps to show that the mass density of stars rises toward the center as $r^{-1.8}$ down to $r = 0.1$ pc to 1.0 pc. The frequent interactions of stars at these short distances leads to their being “thermalized” to an isothermal state just as if they were in a Maxwellian distribution like a stellar gas. However, velocities increase near the center in a way suggesting that considerably more mass is present in or near the center.

Sgr A* and Its Neighbors

Schödel and Genzel et al 2003⁴⁹⁶ have shown stars orbiting at close distances from the galactic center at Sgr A*, particularly star S2, which has an orbital period of only 15.2 years, eccentricity 0.87, a semimajor axis of about 1000 AU, perigalacticon of 1.8×10^{13} m = 120 AU = 17 light-hours. The mass interior to S2 must be about $3.7 \times 10^6 M_\odot$. The Sagittarius A complex is described in *IMA2* p. 927-929. An unusual molecular circumnuclear ring is a doughnut shaped structure with inner radius 2 pc, perhaps a SN remnant with clumpy areas from previous shocks...

Sgr West includes Sgr A*, which appears to coincide with a source of X-rays. Sagittarius A* is a strong radio point source of < 2 AU size. The J2000.0 celestial coordinates of the galactic center (specifically the suspected supermassive black hole Sgr A*) are approximately

$$\begin{aligned}\alpha_{\text{Sgr A}^*} &= 17h\ 45m\ 40s \\ \delta_{\text{Sgr A}^*} &= -29^\circ\ 00'\ 27.9''\end{aligned}$$

which differ slightly from the previously accepted coordinates of the galactic center.

Sgr A* appears to be the location of a supermassive black hole of mass $3.7 \times 10^6 M_\odot$ and a Schwarzschild radius of $16 R_\odot$. (The Schwarzschild radius defines the event horizon.) The UV luminosity is $10^7 L_\odot$, the effective temp 35,000 K. The observed accretion rate is

10^{-3} to $10^{-2} M_\odot \text{ yr}^{-1}$, which could easily provide the observed luminosity. Sgr A* appears to be nearly at rest in the center of the galaxy—it is extremely resistant to movement by gravitational tugs. Strong winds or ejected material are blowing away from Sgr A*. The region must have been violent in the recent past. X-ray flare-ups are observed of 160 times baseline activity currently.⁴⁹⁷ Gamma rays have been detected, including 511 keV photons from positron annihilations and the 1.8 MeV emission of decay of the relatively short-lived isotope ^{26}Al to ^{26}Mg . The region is clearly a dynamic environment. (However, other galaxies such as Andromeda exhibit much more active galactic nuclei.)

⁴⁹⁵ Navarro JF, Frenk CS, & White SDM, “The Structure of Cold Dark Matter Halos”, *ApJ* v.462, p.563

⁴⁹⁶ Eisenhauer F, et al. “A Geometric Determination of the Distance to the Galactic Center” *Astrophysical Journal Letters* Volume 597 Number 2 2003 *ApJ* 597

⁴⁹⁷ Baganoff FK, et al “Chandra X-ray spectroscopic imaging of Sagittarius A* and the central parsec of the Galaxy”, *ApJ* 591:891–915, July 2003, http://docs.lib.psu.edu/physics_articles/368

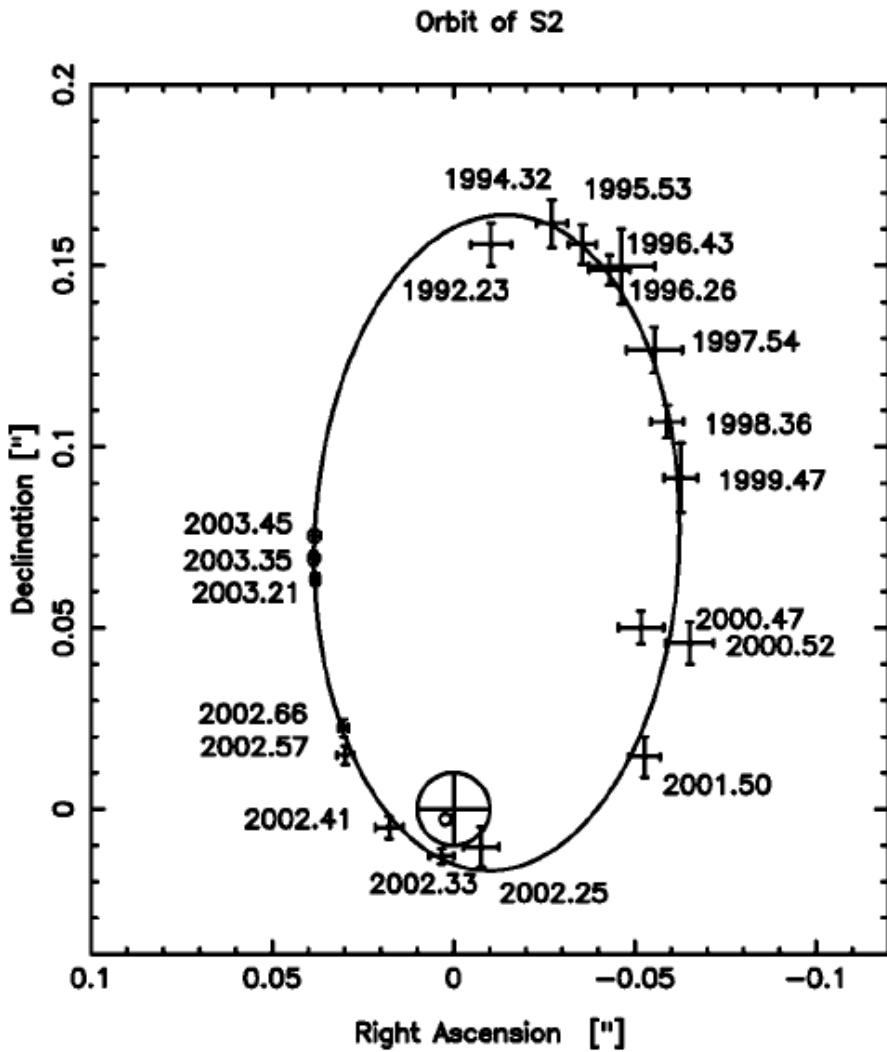


FIG. 1.—Position measurements of S2 in the infrared astrometric frame. Crosses (denoting 1σ error bars) with dates mark the different position measurements of S2, taken with the MPE speckle camera SHARP on the NTT (between 1992 and 2001) and with NACO on the VLT (in 2002 and 2003). The continuous curve shows the best-fit Kepler orbit from Table 1, whose focus is marked as a small error circle. The focus of the ellipse is within a few milliarcseconds at the position of the compact radio source, which is marked by a large circled cross. The size of the cross denotes the ± 10 mas positional uncertainty of the infrared relative to the radio astrometric reference frame.

S2 is a star in a close elliptical orbit about the compact radio source Sgr A* marked by \oplus^{498}

⁴⁹⁸ Eisenhauer F, et al., 2003, op. cit.
Page 188 of 249

The Nature of Galaxies (Chapter 25)

Definition and Scope of Galaxies

A galaxy is a massive, gravitationally bound system that consists of stars and stellar remnants, an interstellar medium of gas and dust, and a poorly understood component called dark matter. Many galaxies, particularly bright ellipticals and spiral galaxies have supermassive black holes at their centers. Sizes of galaxies range from dwarfs with as few as 10^7 stars to giants with as many as 10^{14} stars, each star orbiting the galaxy's center of mass. There are probably more than 170 billion galaxies in the observable universe.⁴⁹⁹

Early Galaxy Discoveries

After early ideas and explorations regarding galaxies of Kant and mapping of the MW by Herschel, systematic cataloging of the galaxies was done by Charles Messier (1730–1817). The first version of Messier's *Catalogue des Nébuleuses & des amas d'Étoiles* contained 45 objects and was published in 1774. The final version of the catalogue was published in 1781—the final list of Messier objects had grown to 103.⁵⁰⁰ On several different occasions between 1921 and 1966, astronomers and historians discovered evidence of another seven deep-sky objects that were observed either by Messier or his friend and assistant, Pierre Méchain, shortly after the final version was published. These seven objects, M104 through M110, are accepted by astronomers as "official" Messier objects.⁵⁰¹ They include gaseous nebulae (including supernova remnants, such as the Crab M1, and giant molecular clouds and H II regions like M16 Eagle and M42 Orion Nebula), planetary nebulae (M27, M57, M76), and globular and open stellar clusters, asterisms (M73), as well as galaxies.

Additional catalogs were made by William Herschel (1738–1822) and his son John. J. L. E. Dreyer (1852–1926) first published the *New General Catalogue of Nebulae and Clusters of Star (NGC)* of 7,840 objects in 1888—most of the Messier objects were included.

William Parsons (1800–1867) first observed spiral structure in galaxies, and Doppler shift confirming rotations was demonstrated first by Vesto M Slipher (1875–1969).

There was great uncertainty about the true nature of many nebulas, culminating in the Great Debate in 1920 of Harlow Shapley (who thought galaxies were local) and Heber D. Curtis (who thought galaxies were extragalactic). Both were drawing on flawed data. Edwin Hubble (1889–1953) settled the debate when he resolved M31 and M33 into individual stars (1923-4) and proved with Cepheid variables that Andromeda M31 was 285 kpc distant (now thought to be 770 kpc), and a galaxy like our own.

The realization that a myriad of galaxies exist other than our own has been of importance comparable to the Copernican Revolution. (IMA2 p. 942)

Fritz Zwicky (1898–1974) first observed in 1933, using the virial theorem, that there is not nearly enough luminous matter to account for observed orbits of cluster galaxies such as the Coma galaxy cluster, and posited unseen dunkle Materie (dark matter).⁵⁰²

Classification of Galaxies

This is reviewed in IMA2 p. 942–948 and only selected comments are included here. Contributions were made by Edwin Hubble, Sidney van den Bergh, Gerard de Vaucouleurs (1918–1995), and the Yerkes (or Morgan) scheme of William Morgan and Philip Keenan.

Hubble introduced in 1926 a graduated sequence (the “tuning fork” diagram) extending (incorrectly) from **early** to **late** types. The earlier types were **ellipticals E0, E1, E2, E3, E4, E5, E6, E7** (which were divided based on numerical degree of apparent **ellipticity** $10\varepsilon = 0\text{--}7$, where **eccentricity** $\varepsilon \equiv 10(1 - (b/a))$). The **transitional lenticulars SO/SBO** could be further subdivided into **S01/SB01, S02/SB02, and S03/SB03**.

⁴⁹⁹ partially paraphrased from <http://en.wikipedia.org/wiki/Galaxy>

⁵⁰⁰ http://en.wikipedia.org/wiki/List_of_Messier_objects

⁵⁰¹ http://en.wikipedia.org/wiki/Charles_Messier

⁵⁰² http://en.wikipedia.org/wiki/Fritz_Zwicky#Dark_matter

The later types were **spirals** (**Sa**, **Sb**, **Sc**, **SBa**, **SBb**, **SBc**), depending on the size of the nucleus and the tightness of the spiral arms) and **Irregulars Ir**. The B in spirals indicated barred (like our Milky Way: SBc). Galaxies were further evaluated for **oblateness** (squashed top to bottom) and **prolateness** (elongated like a football). In reality, galaxies may have 3 orthogonal axes of unequal length and the inclination can be hard to determine, complicating the analysis.

de Vaucoulers introduced variations in classifying galaxies, adding Sd, SBd, Sm, SBm, Im and Ir, and called normal spiral SA, weak bars SAB, and added recognition of dust S0₁, S0₂, S0₃.

The modern sequence is a confused hybrid: **E0**, **E1**, **E2**, **E3**, **E4**, **E5**, **E6**, **E7**, **S0₁**, **S0₂**, **S0₃**, **Sa**, **Sab**, **Sb**, **Sbc**, **Sc**, **Scd**, **Sd**, **Sm**, **Im**, and **Ir**. A similar sequence with Bars B: **SB0₁**, **SB0₂**, **SB0₃**, **SBa**, **SBab**, **SBb**, **SBbc**, **SBc**, **SBcd**, **SBd**, **SBm**. Other taxonomists have added **R** for outer rings, and **s** or **rs** or **r** for no inner ring, weak inner ring, or strong inner ring, respectively.

There is also a **cD**-type galaxy. The "c" in "cD" refers to the fact that the galaxies are very large, hence supergiant, while the "D" refers to the fact that the galaxies appear diffuse. cD's are the largest galaxies and may result from galaxy mergers in galaxy clusters.⁵⁰³

Sidney van den Bergh introduced luminosity classes with Roman numerals **I – V**, and in-between designations are made. For example, the Milky Way is fully classified **SBbcI-II**.

Modern surveys like SDSS utilize automated galaxy classification.⁵⁰⁴

Spiral and Irregular Galaxies

Spiral galaxies each have (1) a disk of stars, gas, and dust rotating about the center, (2) a central stellar bulge resembling an elliptical galaxy, (3) a nearly spherical halo of stars, many in globular clusters, and (4) a supermassive black hole at the very center.⁵⁰⁵

Galaxy brightness may be affected by extinction.

A “**K-correction**” must be used for distant extragalactic objects due to **redshift**. (IMA2 p. 950) “The need for a K-correction arises because an astronomical measurement through a single filter or a single bandpass only sees a fraction of the total spectrum, redshifted into the frame of the observer. So if the observer wants to compare the measurements through a red filter of objects at different redshifts, the observer will have to apply estimates of the K corrections to these measurements to make a comparison...”⁵⁰⁶

The **brightness of the background sky** must also be considered and subtracted from measurements on galaxies. Like other extended objects, the sky brightness μ_{sky} is expressed as **surface brightness**, usually in units of magnitude per square parsec. The darkest sky brightness⁵⁰⁷ is about 22 mag/arcsec² or 220 S₁₀ units (where an S₁₀ unit is defined as the surface brightness of a star whose V-band magnitude is 10 and whose light is smeared over one square degree, or equivalently 27.78 mag arcsec⁻²). Other sites have V-band sky brightness in mag./arcsec² as follows: Mount Wilson 19.8, Palomar 21.5, Haute Provence 21.8, Lowell Mars Hill 20.5, Atacama Paranal⁵⁰⁸ site 22.6 (in blue). Modern CCDs can measure brightness levels as low as 29 mag arcsec⁻². (IMA2 p. 950)

⁵⁰³ cD-type galaxies:

- http://en.wikipedia.org/wiki/Type-cD_galaxy and
- Tonry JL, “Properties of cD galaxies”, *Structure and dynamics of elliptical galaxies*. Reidel Publishing Co., 1987, p. 89-96

⁵⁰⁴ http://www.astro.virginia.edu/class/whittle/astr553/Topic02/Lecture_2.html

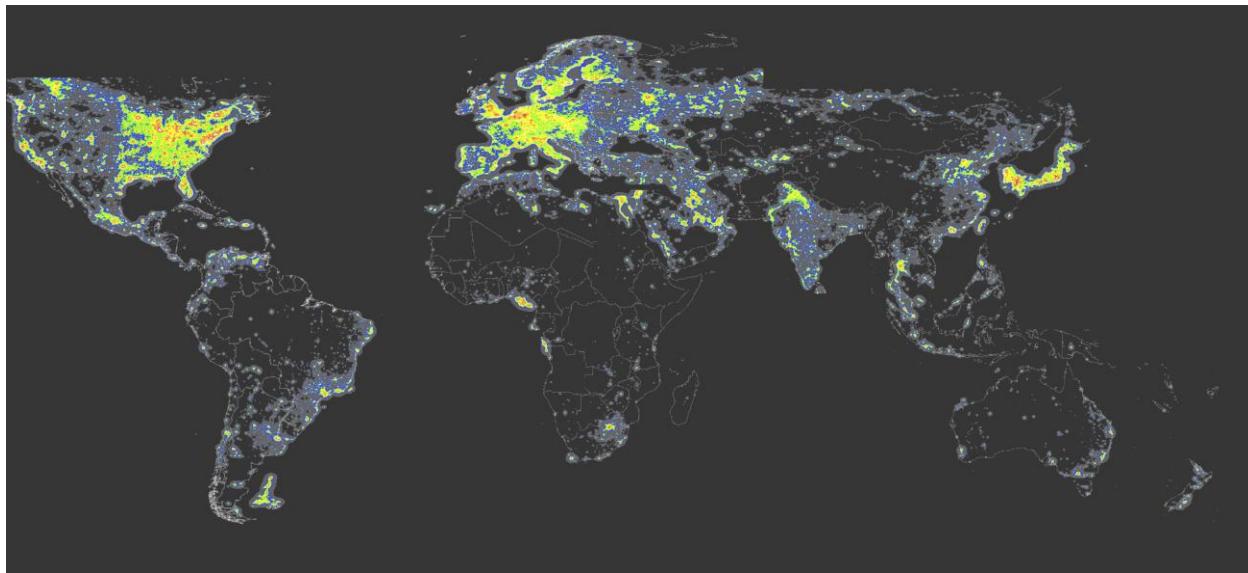
⁵⁰⁵ http://en.wikipedia.org/wiki/Spiral_galaxy

⁵⁰⁶ http://en.wikipedia.org/wiki/K_correction

⁵⁰⁷ Sky Brightness:

- http://en.wikipedia.org/wiki/Sky_brightness
- http://www.astropix.com/HTML/L_STORY/SKYBRITE.HTM

⁵⁰⁸ http://www.eso.org/~fpatat/science/skybright/zenit/zenit_paper.htm



World Atlas of the Artificial Night Sky Brightness⁵⁰⁹

Isophotes are contours of constant surface brightness. These are used to define the radii of galaxies, since there is no clearly defined cutoff. The **Holmberg radius r_H** is a commonly used standard, being the projected length of the semimajor axis of an ellipsoid having an isophotal surface brightness of $\mu_H = 26.5$ B-mag. arcsec $^{-2}$ (i.e., in blue light). Clearly to use a standard this faint in terrestrial astronomy, one must subtract out the sky brightness.

The effective radius r_e is defined as the projected radius within which one-half of the galaxy's light is emitted. The quantity μ_e = surface brightness at the effective radius r_e .

de Vaucouleurs and Sérsic (Sersic) Profiles

A purely empirical relationship between surface brightness for arbitrary r , namely $\mu(r)$, is given for spiral galaxy **bulges** and for **large ellipticals** by an $r^{1/4}$ or **de Vaucouleurs profile** previously mentioned in units of L_\odot per pc 2 , but here written in units of mag arcsec $^{-2}$ (band not stated):

$$\mu(r) = \mu_e + 8.3268 \left[\left(\frac{r}{r_e} \right)^{1/4} - 1 \right]$$

In contrast, **disks** including spiral disks are frequently modeled in surface brightness units of mag arcsec $^{-2}$ by the empirical relationship

$$\mu(r) = \mu_0 + 1.09 \left[\left(\frac{r}{h_r} \right)^{1/1} - 1 \right]$$

where h_r is the characteristic scale length along the midplane and the Sérsic index = 1 (shown here as 1/1), a relationship called “**exponential**”. In general, the exponent takes on the form of $1/n$ for differing **Sérsic Profiles**, where n = “Sérsic index” and controls the degree of curvature of the profile (larger n makes the surface curve more gradually over greater distance). Most galaxies are fit by Sérsic profiles with indices in the range $1/2 < n < 10$. The best-fit value of n correlates with galaxy size and luminosity, such that bigger and brighter galaxies tend to be best fit with larger n .⁵¹⁰

⁵⁰⁹ <http://www.lightpollution.it/worldatlas/pages/fig1.htm>

⁵¹⁰ http://en.wikipedia.org/wiki/Sersic_profile

Rotation Curves of Galaxies

The Sersic profiles alone do not convey information about the distribution of dark matter—this may be gleaned from the galaxy's rotation curve and by other means.

Vera Rubin studied these in the 1970s and reported in 1975 her discovery of the apparently non-Newtonian velocities of stars in spiral galaxies. (See also Differential Galaxy Rotation pertaining to the Milky Way, above.) With her colleagues, she published families of rotation curves for spiral galaxies of types Sa, Sb, and Sc in 1985. These compared the rotational (tangential) linear velocity to the radius, extending from the galactic center to the R_{25} isophotal radius. The R_{25} is the isophotal radius for which blue-band surface brightness has fallen to 25 B-mag/arcsec².⁵¹¹ The flattening of the curves past about $0.25 R_{25}$ to $0.5 R_{25}$ is apparent. (The comparable diameter, $D_{25} = 2R_{25}$, is also used in expressing galaxy size.)

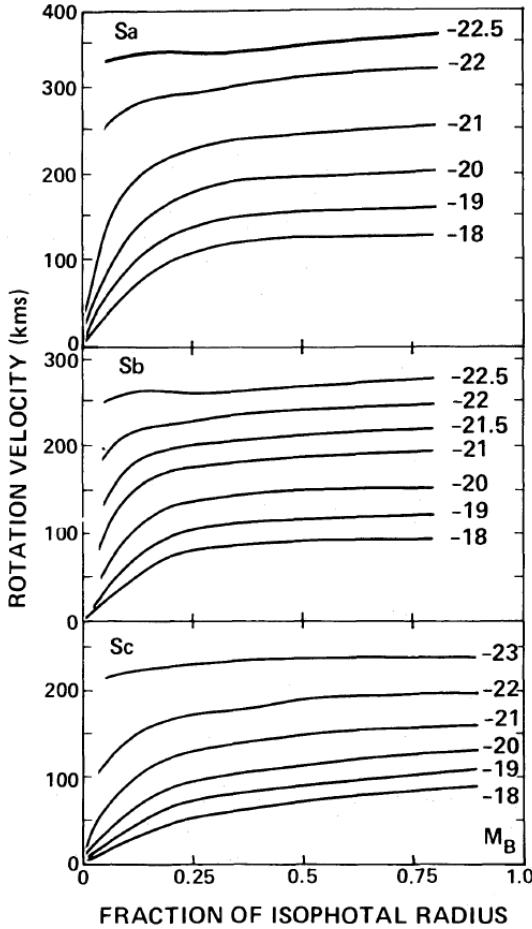


FIG. 4.—Synthetic rotation curves showing average smoothed rotation velocity as a function of fraction of isophotal radius, R_{25} , for (top) Sa galaxies of successive luminosities, (middle) Sb galaxies, and (bottom) Sc galaxies.

For a particular Hubble galaxy type, galaxies that are more luminous—i.e., have more negative M_B —have more rapidly rising and larger final values of V_{\max} . For the same luminosity, early Sa galaxies have larger V_{\max} than late Sc galaxies. In general, the rotation curves suggest that the gravitational potentials are not following the distribution of known luminous matter, thought to be a signature for dark matter (IMA2 p. 951)

The V_{\max} rotation speed of irregular galaxies is only 50 to 70 km s⁻¹, perhaps indicating a lower threshold limit on V_{\max} below which well-formed spiral galaxies do not develop.

See also discussion under Ellipticals: Shape and Rotation regarding rotationally supported shapes versus pressure-supported shapes.

⁵¹¹ Rubin VC, et al, “Rotation velocities of 16 SA galaxies and a comparison of Sa, Sb, and SC rotation properties”, *ApJ*, Part 1, vol. 289, Feb. 1, 1985, p. 81-98, p. 101-104.

The Tully Fisher Relation

These empirical formulas express a roughly linear relationship between luminosities of spiral galaxies in various wavelength bands versus their V_{\max} . It was developed by R. Brent Tully and J. Richard Fisher and published in 1977. It has been applied to 21-cm radio H I emissions (which exhibit double peaks from the Doppler shifted galactic periphery approaching and receding, shown below), blue-band visible light M_B (good for young hot stars), and infrared wavelengths such as M_{H^i} (good for overall distributions of luminous mass). (IMA2 p. 952-6). Distance measurement discussed later.

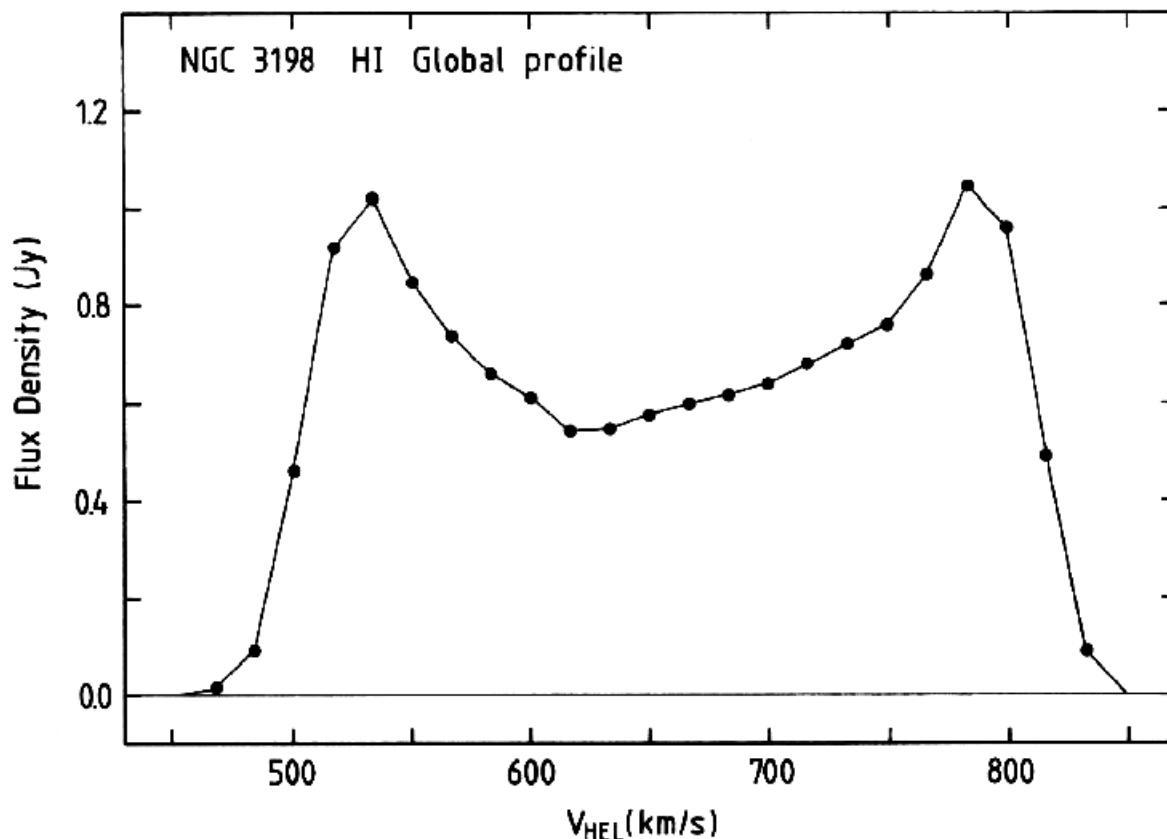


Fig. 2. Global H I line profile. The flux densities have been corrected for primary beam attenuation

21-cm radio H I emissions from galaxy NGC 3198, which exhibit double peaks from the Doppler shifted galactic periphery segments approaching toward and receding from the observer. Spectral Flux density in Janskys.

V_{HEL} is heliocentric (Sun-based) velocity in km s^{-1} , averaging 660 km s^{-1} recession(Begeman 1989). ⁵¹²

The Radius-Luminosity Relation

For early spiral Sa-Sc galaxies of increasing luminosity, their isophotal R_{25} radii (measured in kpc) increase according to an empirical relationship. (IMA2 p. 956):

$$\log_{10} R_{25} = -0.249M_B - 4.00$$

⁵¹² Begeman KG, "H I rotation curves of spiral galaxies. I - NGC 3198", *Astronomy and Astrophysics*, vol. 223, no. 1-2, Oct. 1989, p. 47-60

Mass and Mass to Light Ratios

Sc's are slightly less massive than Sa's. However, the average ratios of M/L_B (expressed in units of M_⊙ / L_⊙^{Bol}) are more strongly dependent on Hubble type, being M/L_B 6.2 for Sa's, 4.5 for Sb's , and 2.6 for Sc's.

Color, Gas, and Dust

The trend of Mass to Light Ratios suggests a greater fraction of massive main sequence “early” stars in “late” Sc's compared to “early” Sa's (not inconsistent and unfortunate Hubble labels), and therefore a bluer color index. Sc's also have greater gas and dust and are more centrally condensed.

Irregulars (such as the LMC) are the bluest of galaxies, suggesting active star formation.

Stars in spiral galaxies are redder in their bulges than in their disks, which relates to metallicity and the relative proportion of star formation. (*IMA2* p. 957-8)

X-ray Luminosity

There is a close linear correlation between X-ray and B-band luminosity, suggesting that X-rays are generated by a small but constant fraction of objects (such as SNRs) in a spiral galaxy.

Supermassive Black Holes SMBH

There is strong evidence of a SMBH at the center of Andromeda M31, where a large non-luminous mass is predicted with mass = 1.4x10⁸ M_⊙. Another SMBH with a mass of 3.2x10⁹ M_⊙ has been inferred in M87, a giant elliptical galaxy.

Black holes mass can also be deduced by velocity dispersion via the virial theorem (*IMA2* p. 959-60), using the derived formula

$$M_{\text{virial}} \cong \frac{5R\sigma_r^2}{G}$$

where the sigma term is the velocity dispersion in the radial velocity and R is the radius of a suitably chosen sphere within the black hole's influence.

Another empirical relation with dispersion is given by $M_{\text{bh}} = \alpha(\sigma/\sigma_0)^\beta$.

It is quite possible that a SMBH is at the heart of most galaxies.

Globular Star Clusters

To compare the frequency of globular clusters (GCs) in galaxies, the quantity **specific frequency S_N** of GCs is used. This term that has been normalized to a standard absolute magnitude of the parent galaxy. S_N is greater in late-type galaxies and in galaxies with higher luminosity. In particular, elliptical galaxies and especially the gigantic cD's have more GCs per unit luminosity.

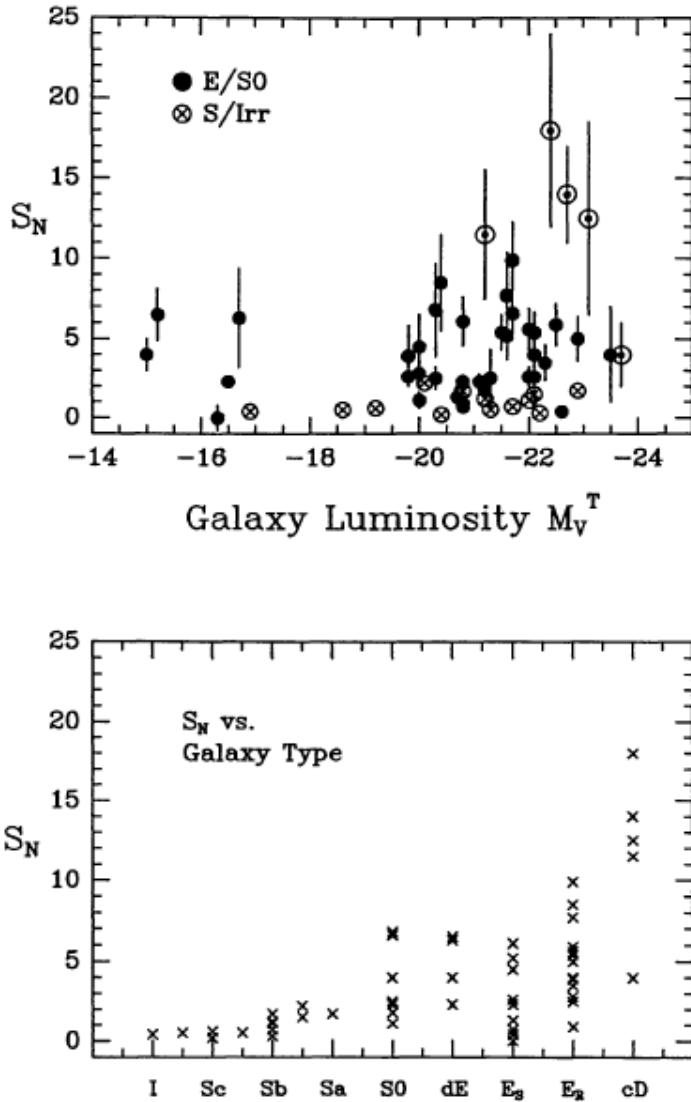


Figure 4 Specific frequency S_N of globular cluster systems, as listed in Table 1. In the upper panel, S_N is plotted versus the luminosity of the parent galaxy; E/S0 types are denoted by the filled symbols, and spiral/irregular types by the crosses. The five giant ellipticals at the centers of rich clusters (Virgo, Fornax, Hydra, Coma, A2199) are denoted by the circled dots. In the lower panel, S_N is plotted against morphological type. Here E_S and E_R refer to ellipticals in sparse and rich clusters, and the last bin (cD) again denotes the five central giant ellipticals.

Specific frequency of globular clusters,
evaluated by luminosity and galaxy Hubble type (Harris 1991⁵¹³)

Spiral Structure

There is a rich and elegant body of mathematical theory which I will report only selectively here. Spiral galaxies may have a **grand design** (such as M51 shown at the beginning of this document) with two very symmetric and well-defined arms, or may have **multiple arms**, or at the other extreme may be poorly defined **flocculent spirals** with only partial arm fragments.

⁵¹³ Harris WE, “Globular cluster systems in galaxies beyond the Local Group”, *Annual review of astronomy and astrophysics*. Vol. 29 (A92-18081 05-90,. 1991

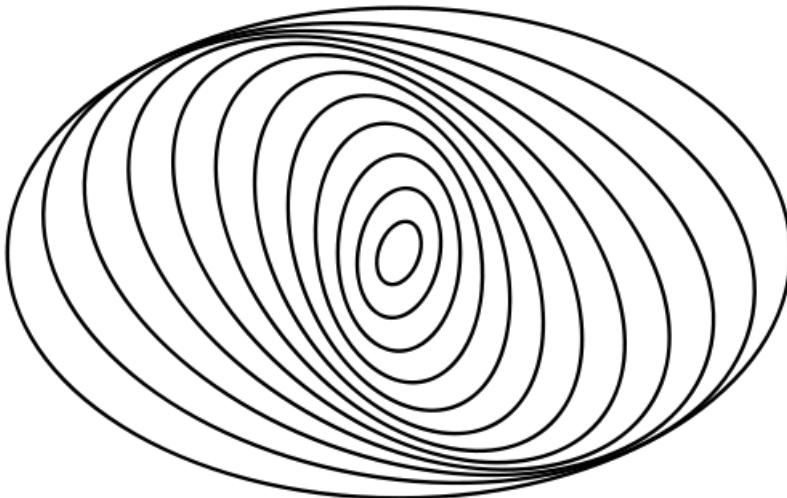
The arms are especially apparent and most distinct in blue band, due to the presence of hot young O and B blue stars, consistent with active star formation. In infrared or red light, the arms appear broader and less pronounced, suggesting that the disk is dominated by older redder stars. Dust and gas preferentially resides on the concave sides of the arms. In general, the spiral bands appear quite different when viewed in varying wavelengths, such as X-ray, UV, IR, and radio.

The overall [?number] density of stars is increased only by about 10 to 20% in spiral arms (ZXI lecture 4).

The convex side of the arm might be expected to be the forward edge (the tip curves back behind, pointing in a direction essentially opposite to the direction of rotation). The concave inner side is the trailing edge. This type of rotation is called a trailing arm structure. However, this is not always the case, the direction of rotation can be hard to determine, and some spiral arms are leading structures. Spiral arms that can be evaluated are usually trailing, however.

Spiral arms **winding** might be expected if they were “material” structures composed of a fixed set of differentially rotating stars, but spiral arms probably do not **wind up**.

Instead, the best hypothesis explaining spiral arms is the **Lin-Shu Density wave theory** (of C. C. Lin and Frank Shu, devised in 1964.⁵¹⁴) The arms are **quasi-static density waves** (which are similar to a local traffic piling up on a freeway, in which individual cars slow and accumulate in the congested area while passing through). In a non-inertial rotating frame with angular **global patterns speed** Ω_{gp} , they appear to be stationary. Stars, dust, and gas move through these regions of higher density (perhaps 10% or 20% higher in density). At the **corotation radius** R_c , the spiral arm and the objects at the same radius move together. Objects closer to the center ($R < R_c$) rotate in the non-inertial frame with shorter periods than the arm, whereas objects further away ($R > R_c$) rotate more slowly than the density wave pattern. Most new star formation occurs closer in to the center than R_c . Density wave formation can cause a 3 or 4 armed spiral, and there are numerous other variants.



A pileup of nested ellipses rotating at slightly different angular velocities may explain a density wave having the shape of spiral arms with a central bar.⁵¹⁵

It is also possible to have mathematically defined **resonances** leading to denser regions where stars accumulate:

- 0 to 2 **inner Lindblad resonances (ILR)**, which may cause **inner rings**,
- an **ultraharmonic resonance**, which may cause an **inner ring**,
- a **corotation resonance (CR)**, which may bound bars, and
- an **outer Lindblad resonance (OLR)**, which may cause **outer rings**.

Individual stars rotating about the galactic midplane exhibit motion that to a first order approximation “can be imagined as being the combination of a retrograde orbit about an epicycle and the prograde orbit of the center of the epicycle about a perfect circle (IMA2 p. 975).

A companion galaxy may be needed to cause the initial tidal interactions triggering spiral arm formation.

⁵¹⁴ http://en.wikipedia.org/wiki/Density_wave_theory

⁵¹⁵ ibid.

Impressive spiral arms and bars have been demonstrated in N-Body simulations.⁵¹⁶

In the case of flocculent spirals, an alternate theory of **stochastic self-propagating star formation** has been advanced by Mueller and Arnett, 1976. The theory posits outbursts of star formation that propagate across the galaxy.

“The number and how tightly the spiral arms are wound are well correlated with other, large scale properties of the galaxies, such as the luminosity of the bulge relative to the disk and the amount of gas in the galaxy. This suggests that there are global physical processes involved in spiral arms.” (ZXI lecture 3).

Elliptical and SO Galaxies

These are remarkably diverse and complex. An elliptical galaxy is a galaxy having an approximately ellipsoidal shape and a smooth, nearly featureless brightness profile. They range in shape from nearly spherical to highly flat and in size from hundreds of millions to over one trillion stars.⁵¹⁷

They tend to have older (and thus reddish) lower-mass stars, with a majority having sparse ISM and dust and minimal star formation activity, with few open star clusters. Unlike spirals, many are dominated not by rotation but by random stellar motions or “pressure”. Rather than Hubble’s system of categorizing ellipticals based on ellipticity (0 - 7), it has been found more useful to categorize them by other factors such as size, absolute magnitude, and surface brightness (*IMA2* p. 983).

⁵¹⁶ Hohl F, “Numerical Experiments with a Disk of Stars” *ApJ* 168 p. 343, 1971

⁵¹⁷ http://en.wikipedia.org/wiki/Elliptical_galaxy, edited



Lenticular S0 Galaxy NGC 4150, mass~ $6 \times 10^9 M_{\odot}$, distance ~13.6 Mpc, $M_B = -18.5$ mag
The inset core width is ~500 pc. It shows a slightly rotated “magnified view of the chaotic activity inside the galaxy's core. The blue areas indicate a flurry of recent star birth.

The stellar breeding ground is about 1,300 light-years across. The stars in this area are less than a billion years old.” Dark strands of dust are also apparent, tentative evidence of a recent merger. (HST images, Oct. and Nov. 2009)⁵¹⁸

Morphological Classes of Elliptical and SO Galaxies

They include:

cD Galaxies:

cD's are immense, rare, up to 1 Mpc in size, found near center of large dense galactic clusters. They have M_B magnitudes of -22 to -25, masses of 10^{13} to $10^{14} M_{\odot}$, and D_{25} diameters of 300-1000 kpc. They have high central surface brightness ($\mu = 18$ B-mag arcsec $^{-2}$) and very extended diffuse envelopes (for which $\mu = 26 - 27$ B-mag arcsec $^{-2}$). They have large numbers of globular clusters and a high specific frequency for globular clusters (averaging $S_N = 15$). They have high mass to luminosity ratios ($M/L_B > 100 M_{\odot}/L_{\odot}$, as high as $750 M_{\odot}/L_{\odot}$), indicating large amounts of dark matter. Surface brightness profiles follow the $r^{1/4}$ law.

⁵¹⁸ Lenticular S0 Galaxy NGC 4150:

- <http://hubblesite.org/newscenter/archive/releases/galaxy/elliptical/2010/38/>
- <http://hubblesite.org/pubinfo/pdf/2010/38/pdf.pdf>

Normal Elliptical Galaxies:

These include giant ellipticals **gE**'s, intermediate luminosity ellipticals **E**'s, and compact ellipticals **cE**'s. E's have M_B magnitudes of -15 to -23, masses of 10^8 to $10^{13} M_\odot$, and D_{25} diameters of 1-200 kpc. They have relatively high central surface brightness. They have specific frequency for globular clusters $S_N = 5$. They have high mass to luminosity ratios ($M/L_B = 7$ to $100 M_\odot/L_\odot$). Hubble's **Lenticular SO**'s and **SBO**'s are often grouped here as well, and have M_B magnitudes of -17 to -22, masses of 10^{10} to $10^{12} M_\odot$, and D_{25} diameters of 10-100 kpc, $M/L_B \sim 10 M_\odot/L_\odot$, specific frequency for globular clusters $S_N \sim 5$. Surface brightness profiles follow the $r^{1/4}$ law (such as shown on *IMA2* p. 986 for NGC 3379)

Dwarf Elliptical Galaxies dE's:

dE's have M_B magnitudes of -13 to -19, masses of 10^7 to $10^9 M_\odot$, and D_{25} diameters of 1-10 kpc. They have lower surface brightness than cE's of the same magnitude. They have specific frequency for globular clusters $S_N \sim 4.8$, still higher than spirals. They have mass to luminosity ratios $M/L_B = \sim 10 M_\odot/L_\odot$. Metallicities are lower than for E's. Surface brightness profiles are "exponential"—they follow the $r^{1/1}$ law, μ has a linear relation to r/h_r , and $I(R) \propto e^{-kR}$.

Dwarf Spheroidal Galaxies dSph's

dSph's are extremely low luminosity low surface brightness objects detected only in the vicinity of the MW. They have M_B magnitudes of -8 to -15, masses of 10^7 to $10^8 M_\odot$, and D_{25} diameters of 0.1-0.5 kpc. They have mass to luminosity ratios $M/L_B = 5 - 100 M_\odot/L_\odot$. No clusters are mentioned. Surface brightness profiles are "exponential".

Blue Compact Dwarf Galaxies BCD's:

BCD's are small compact galaxies that are unusually blue, with color index $B - V = 0.0$ to 0.3, thus like spectral class A stars, suggesting they are undergoing active star formation. They have M_B magnitudes of -14 to -17, masses of $\sim 10^9 M_\odot$, and D_{25} diameters of < 3 kpc. They have low mass to luminosity ratios $M/L_B = 0.1 - 10 M_\odot/L_\odot$. They have an abundance of H I and H II. No clusters are mentioned.

Gravitational Binding, Dust and Gas, and Star Formation

Ellipticals (dE's and dSph's) tend to have low gravitational binding and retain only small amounts of gas and dust (esp. dSph's). Therefore they are not now actively forming stars. The dust sometimes counterrotates with respect to the stellar components. Perhaps gas has been stripped by supernovas or galaxy interactions. What interstellar medium is present tends to be hot, with $T > 10^6$ K (*ZXI* lecture 5).

Metallicity

Ellipticals have low metallicities, such as seen with globular clusters. Metallicity correlates with luminosity—metallicity of dwarf elliptical and ellipticals tends to increase with absolute blue-magnitude M_B (*IMA2* p. 958). As with spirals, metallicity gradients are seen with ellipticals...

The Faber-Jackson Relation

This expresses, for dE's, dSph's, and E's as well as for spiral bulges, a correlation between central radial velocity dispersion σ_0 and the luminosity corresponding to M_B , namely

$$L_B \propto \sigma_0^\alpha$$

where a was originally taken to be 4 by Faber and Jackson,⁵¹⁹ but is now fitted to a range of values. *IMA2* p. 988 describes further efforts to fit the observations to an empirical equation relating luminosity, velocity dispersion, and effective radius r_e . The relation exhibits considerable scatter for ellipticals.

Shape and Rotation

The source of the shape of ellipticals is somewhat unclear. Most have no preferred axis of rotation. It is likely that some of the gas, dust, globular clusters, and even dwarf galaxies have been captured since the galaxies formation. Their shapes are not generally due to rotation, as the velocity dispersions often exceed the rotational velocities. The shapes may be due to anisotropic velocity dispersions.

As an extreme example, the E3-E4 galaxy NGC 1600 has $V_{\text{rot}} = 1.9 \text{ km s}^{-1}$ and an extremely low $V_{\text{rot}}/\sigma = 0.013$.

The rotation parameter $(V/\sigma)^*$ is defined as

$$\left(\frac{V}{\sigma}\right)^* \equiv \frac{(V_{\text{rot}}/\sigma)_{\text{observed}}}{(V_{\text{rot}}/\sigma)_{\text{isotropic}}}$$

Here, the isotropic term in the denominator is “the value expected for an isotropic oblate spheroid flattened purely by rotation”.⁵²⁰ A galaxy is considered to be **rotationally supported** if rotation parameter $(V/\sigma)^* \geq 0.7$. However, it is < 0.016 for this galaxy, thus it is not rotationally supported. NGC 1600 has no statistically significant axis of rotation (*IMA2* p. 989).

Most bright E’s and gE’s have $(V/\sigma)^* \sim 0.4$, and are considered to be **pressure-supported**, meaning that their shapes are due to random stellar motions as with an ideal gas, rather than to overall rotation.

However, less luminous galaxies for which $-18 > M_B > -20.5$ (including cE’s) typically have rotation parameters $(V/\sigma)^* \approx 0.9$ and are therefore rotationally supported. (*IMA2* p. 989) (Spiral bulges are also generally rotationally supported.)

The shapes of these elliptical galaxy isophotal surfaces are subtly “boxy” or “disky”. In the Fourier expansion of the polar coordinates representation of shape, given by radius $a(\theta)$ of this surface, the term $a_4 \cos(4\theta)$ is associated with the boxy/disky shape component. If $a_4 < 0$, the surface is boxy (and these tend to be pressure supported, etc.), whereas if $a_4 > 0$, the surface is disky (and these tend to be rotationally supported, etc.) Although there are many factors involved, boxy galaxies may reflect a signature of past galaxy mergers or tidal interactions (*IMA2* p. 991).

Relative Numbers of Galaxies by Hubble Type

In order to express the relative number of galaxies by Hubble type, it is useful to define the **luminosity function LF** expressed as $\phi(M) dM$, which gives the number of galaxies in a sample with absolute magnitudes between M and $M+dM$. $\phi(M)$ has units of galaxies per magnitude per pc³. The following graph (Binggeli, Sandage, Tammann 1988)⁵²¹ gives the log of luminosity functions $\phi(M)$ for a selection of 121 galaxies near the MW (“local field”) and also for a much larger sample of 1300 galaxies in the Virgo cluster (which is considered much more reliable). M_{BT} is absolute blue-band total magnitude (i.e., not surface brightness), and “Total” is the sum of the individual luminosity functions (including types that are not specifically graphed; see also *IMA2* p. 991). Galaxies of type dE and dwarf Irr are the greatest in numbers in Virgo, but S’s and E’s are the most prominent in cluster luminosity and mass. [The latter statement in *IMA2* is not quite so obvious to me, but presumably true due to the much higher magnitudes of the S and E galaxies]. The ellipticals are more

⁵¹⁹ Faber SM, Jackson RE, “Velocity dispersions and mass-to-light ratios for elliptical galaxies”, *Astrophysical Journal*, vol. 204, Mar. 15, 1976, pt. 1, p. 668-683

⁵²⁰ Houjun Mo, Frank Van Den Bosch, Simon D. M. White, *Galaxy Formation and Evolution*, Cambridge UP, 2010, p. 47.

⁵²¹ Binggeli B; Sandage A, Tammann GA, “The luminosity function of galaxies”, *Annual review of astronomy and astrophysics*. Vol. 26, 1988, p. 509-560.

prominent in Virgo than the local field, and BCDs are only seen in Virgo. The Virgo values reported (Binggeli p. 550) are E's=12%, SO's=26% and S+Irr's=62%, whereas the values for the Coma cluster (same page) differ significantly. Clearly the relative proportions of galaxy types depends on the population sampled.

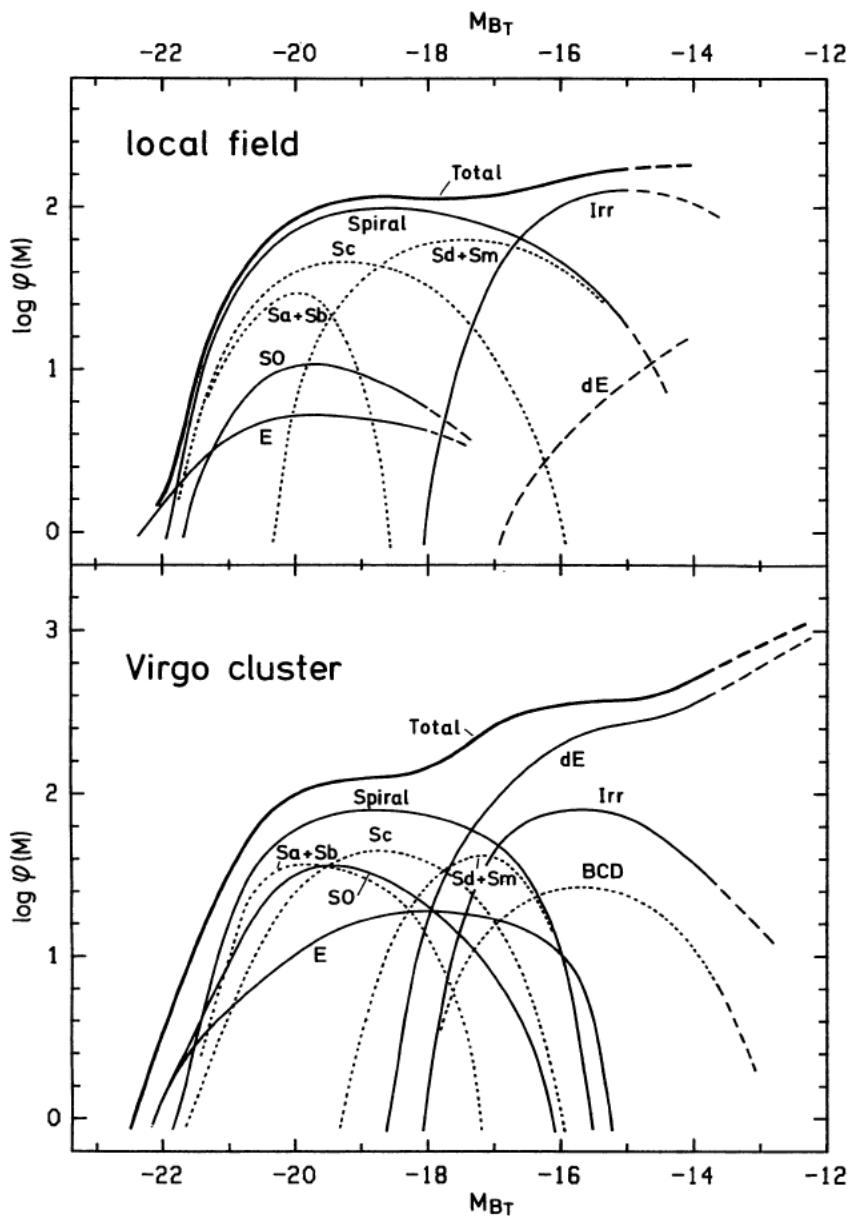


Figure 1 The LF of field galaxies (top) and Virgo cluster members (bottom). The zero point of $\log \varphi(M)$ is arbitrary. The LFs for individual galaxy types are shown. Extrapolations are marked by dashed lines. In addition to the LF of all spirals, the LFs of the subtypes Sa + Sb, Sc, and Sd + Sm are also shown as dotted curves. The LF of Irr galaxies comprises the Im and BCD galaxies; in the case of the Virgo cluster, the BCDs are also shown separately. The classes dS0 and "dE or Im" are not illustrated. They are, however, included in the total LF over all types (heavy line).

Graph of Binggeli, Sandage, Tammann 1988.⁵²²

A luminosity function $\Psi(M,z)$ can also be derived, a function of both luminosity and redshift distance z , thus in the luminosity-“position” plane.

⁵²² ibid.

ZXI (lecture 6) showed a graph from an article he co-authored in which blue galaxies have a different LF from red galaxies.⁵²³

The following graph of SDSS data plots a luminosity function $\Phi(M, z)$ against M_r for $z=0.1$, which is near the median redshift of the sample. I have not mastered this complex subject matter. The authors measured the galaxy luminosity density at $z = 0.1$ in five optical bandpasses corresponding to the SDSS bandpasses shifted to match their rest-frame shape at $z = 0.1$. These included $^{0.1}r$, for which $T_{\text{eff}} = 5595$ K for the r bandpass. In this case, Φ is not decomposable into a product of separate distributions $f(M)$ and $n(z)$. The relative peaks on the left found in small numbers represent the largest ellipticals, including cD galaxies.⁵²⁴

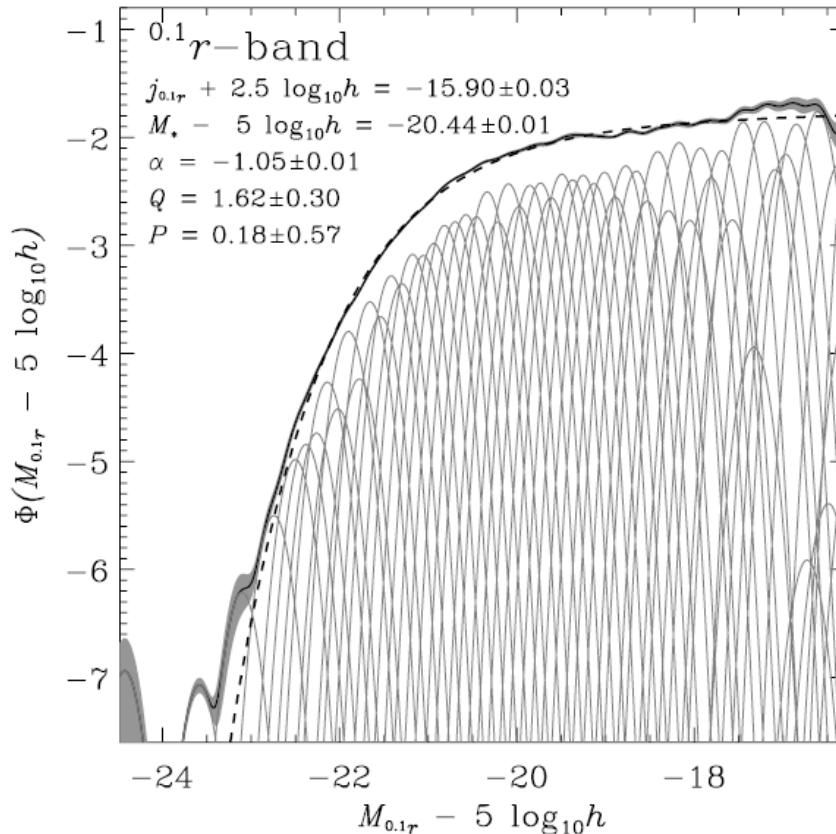


FIG. 5.—Luminosity function in the $^{0.1}r$ band. The thick solid line is the luminosity function fit; the thin solid lines are the individual Gaussians of which it is composed. The gray region around the luminosity function fit represents the 1σ uncertainties around the line; naturally, these uncertainties are highly correlated with each other. The dashed line is the Schechter function fit to the result. The luminosity density, the evolution parameters, and the parameters of the Schechter function are listed in the figure.

⁵²³ Baldry IK et al, “Quantifying the Bimodal Color-Magnitude Distribution of Galaxies”, *The Astrophysical Journal*, Volume 600, Issue 2, pp. 681-694, 01/2004. Figure 7.

⁵²⁴ Blanton MR et al, “The Galaxy Luminosity Function and Luminosity Density at Redshift $z = 0.1$ ”, *The Astrophysical Journal* Volume 592 Number 2 2003.

Galactic Evolution (Chapter 26)

Interaction of Galaxies

Most galaxies have not arisen in isolation but in groups or clusters (*IMA2* p. 999), and their interactions play an important role in their evolution. Intergalactic spacing is typically only 100 times their diameters, and some clusters (such as Coma Cluster) are more densely populated. Ellipticals are more common in the center areas than the outer regions of such galaxy clusters, especially for well-formed clusters. These observations may correlate with an increased probability of galaxy interactions and mergers, leading to increased velocity dispersion and a $r^{1/4}$ de Vaucouleurs profile.

Many disk galaxies have warped disks in H I emission.

The gas in galaxies in a cluster appears to be pulled out of the galaxies into the surrounding intergalactic space in the cluster, where it is seen to be hot and X-ray emitting.

Dynamical Friction

The individual stars in colliding galaxies rarely physically collide, due to their small size relative to the interstellar spacing, but galaxies themselves do commonly collide or pass through each other. **Dynamical friction** describes the loss of momentum and kinetic energy of moving bodies arising through an indirect gravitational interaction with the surrounding matter in space. It is sometimes referred to as gravitational drag, and was first discussed in detail by Chandrasekhar in 1943.⁵²⁵

This quasi-frictional force takes the form of

$$f_d \cong C \frac{G^2 M^2 \rho}{v_M^2}$$

where in an idealized example

M is the mass of a smallish approaching object Obj1 (perhaps a globular cluster or small galaxy)

v_M is the average velocity of Obj1 relative to Obj2,

ρ is the density of Obj2, where Obj2 is a very extensive object (such as a large galaxy),

C is a function that depends on how v_M compares to the velocity distribution σ of Obj2, and

f_d is the resultant drag force on Obj1.

A **wake trail** is left behind Obj1 as it passes through Obj2, and some of Obj1's kinetic energy is transferred to the stars of Obj2. (MCM: I am assuming that the stars in this trail have arisen from both Obj1 and Obj2)

The frictional force is stronger for greater ρ . It is also apparent that slow encounters with small v_M exert much more dynamical frictional drag on Obj1 than fast fly-throughs, for which there is less time for interaction.

The dynamical friction force exerts a torque on clusters, reducing their angular momentum, causing many of them eventually to spiral in to the center of their galaxy and disappear, thus depleting the galaxy of star clusters, especially the lower mass clusters. (Details of the explanation are somewhat unclear to me, see *IMA2* p. 1001-4 and *here*⁵²⁶). Even satellite galaxies are affected, explaining why it is believed that the Magellanic clouds will someday merge with the MW.

Frictional effects from a collision with a galaxy in a retrograde orbit may explain why a galaxy can end up with a counterrotating core.

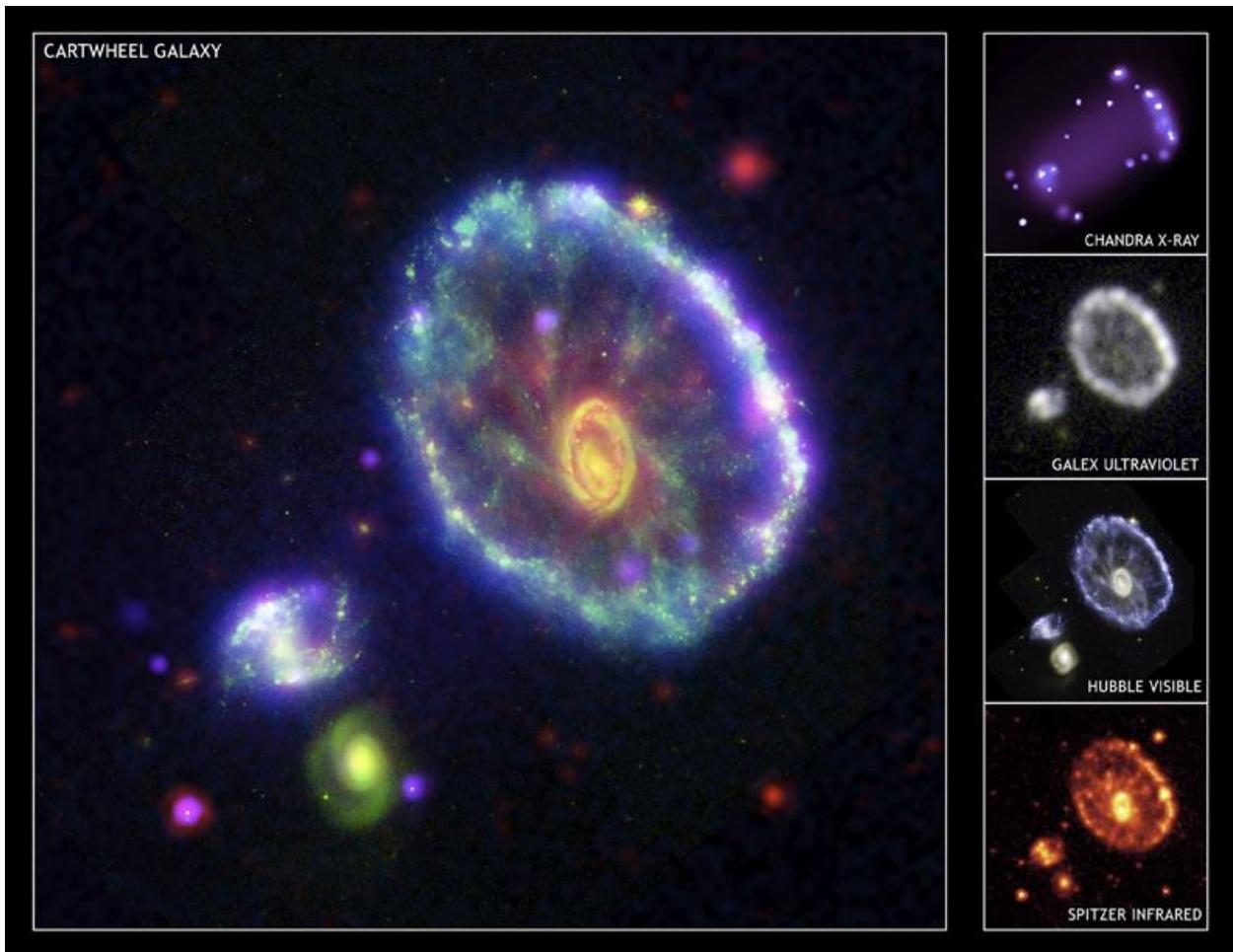
⁵²⁵ http://en.wikipedia.org/wiki/Dynamical_friction

⁵²⁶ <http://www.research-in-germany.de/93492/2012-02-15-how-globular-star-clusters-survive-collisions.html>

Rapid Encounters, Rings, Tidal stripping, and Streams

Collisions at fast v_M 's cause an increase of KE in the collided galaxies. In order for the galaxies to return to virial equilibrium, the added KE must be adjusted for, either by increasing star spacing, or by having it carried away from the galaxy, often by the most energetic components in the form of a **stream** of gas and stars (cf. the Magellanic Stream). This "evaporation" cools the galaxy. "[the Magellanic stream began] at about the time when the two Magellanic Clouds may have passed close to each other, triggering massive bursts of star formation.... The strong stellar winds and supernova explosions from that burst of star formation could have blown out the gas and started it flowing toward the Milky Way"⁵²⁷

High speed nearly head-on galactic collisions may also result in a **ring galaxy** such as the Cartwheel (ESO 350-40), in which the ring represents a sudden brief burst of star formation:



Cartwheel Galaxy (ESO 350-40), composite image in false colors, using

- (1) Galaxy Evolution Explorer in ultraviolet light (blue); (2) HST in visible light (green);
- (3) the Spitzer Space Telescope in IR (red); and (4) Chandra X-ray Observatory (purple).

The Cartwheel galaxy is one of the brightest ultraviolet energy sources in the local universe.

In some visible-light images, it appears to have spokes."

Distance 500 Mly = 150 Mpc, ~1 arcmin (NASA etc.)⁵²⁸

Gaseous rings may be formed by collisions—there are no stars in these rings (see for instance M105).

Galaxies that are bound to each other eventually merge.

⁵²⁷ <http://www.astronomynow.com/news/n1001/05streamer/>

⁵²⁸ Cartwheel Galaxy (ESO 350-40):

- <http://www.nasa.gov/centers/jpl/news/galex-20060111.html>
- http://en.wikipedia.org/wiki/Cartwheel_Galaxy

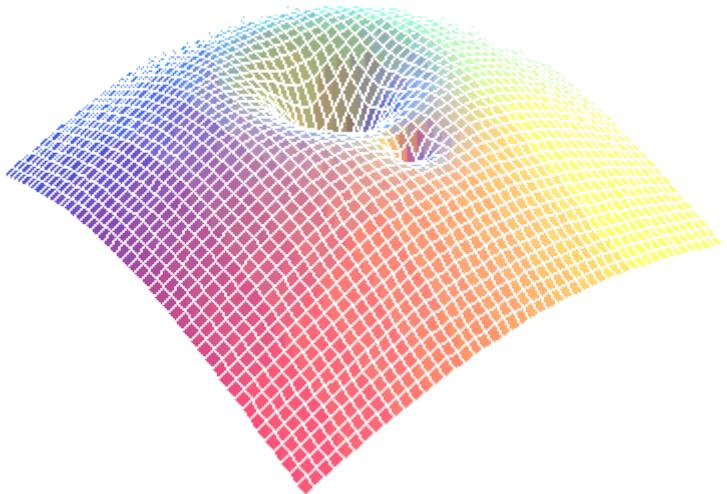
Tidal stripping can dissipate the orbital energy by pulling stars and gas out of one or both galaxies, such as caused the Magellanic Stream. This occurs when material extends beyond (“overflows” or escapes from the potential well of) the retaining equipotential lines. These potential barriers are analogous to Roche lobes for stars.

The diagram⁵²⁹ to the right depicts a gravitational equipotential surface, showing potential wells bounded by Roche Lobes (not explicitly delineated) that are separated by the L₁ Lagrangian point, here for close binary stars and by analogy for galaxies. (See full discussion of a similar diagram in Chapter 18)

The tidal radii l_1 and l_2 define the respective distances to the L₁ Lagrangian point (the point where escape to the other object is the easiest, see diagram below) for the two colliding objects. (IMA2 p. 1006, 657, 659) Tidal stripping may also explain the retrograde orbits of about one half of the stars in the outer stellar halo of the Milky Way (IMA2 p. 1009)

Collisions may also result in **tidal tails**, such as are dramatically evident in the well-known colliding Antennae galaxies NGC 4038 and 4039.⁵³⁰ According to simulations, tails are usually seen only with close slow encounters, which produce tidal deformations in the equipotential surfaces analogous to Earth’s tidal bulging. The stars and gas torn from the near side of a tidally bulging colliding galaxy form an apparent **bridge** to the other galaxy (often obscured, may not actually connect in an optimal view), whereas the material bulging and escaping on the opposite side will form a tail that conserves angular momentum (IMA2 p. 1009 and *here*⁵³¹).

Galaxies have been found that have **polar rings**⁵³² (usually around ellipticals or SO galaxies, such as NGC 4650A) or are **dust-lane ellipticals** (such as NGC 5128), in either case exhibiting material stripped from colliding galaxies.



Modeling Interactions With N-Body Simulations

This approach has become an indispensable but complex and costly component of exploring galactic interactions. It was pioneered by Alar and Juri Toomre in 1972. Modern simulations must include dark matter and self-gravitation. Examples of simulations are available on the Web,⁵³³ including links at the beginning of this document. Simulations have helped for instance to show why outer clusters of the Milky Way were probably formed elsewhere and captured later. Satellite galaxy collisions may explain the warps that persist for billions of years in some galaxy disks. (Details are mostly omitted here. See IMA2 p. 1008-1011, etc.)

Starburst Galaxies

These are often colliding and interacting galaxies emitting blue color from prominent hot new OB star formation. They have increased luminosity (often obscured by dust in visible and UV bands, but bright and well seen in IR). The star formation may be very rapid, using up to 300 M_\odot per year, but may only last 20 million years. Star formation is not necessarily concentrated in the galactic center, and may extend over much of the disk. M82 is an example:

⁵²⁹ <http://csep10.phys.utk.edu/astr162/lect/binaries/accreting.html>

⁵³⁰ http://en.wikipedia.org/wiki/Antennae_Galaxies

⁵³¹ http://en.wikipedia.org/wiki/Galactic_tide

⁵³² http://en.wikipedia.org/wiki/Polar-ring_galaxy

⁵³³ <http://hubblesite.org/newscenter/archive/releases/2008/16/video/d/> etc.



M82 (NGC 3034) Starburst Galaxy by HST, 2006

Galaxy at distance 3.7 Mpc, width 8 arcmin or ~8.5 kpc (thus smaller than the MW)

Composite of Multiple filters: F658N (Hydrogen-alpha) assigned red/orange; F814W (I) assigned red
F555W (V) assigned green; and F435W (B) assigned blue⁵³⁴

According to some definitions, starburst galaxies include Blue compact galaxies (BCGs), Luminous infrared galaxies (LIRGs) and Ultra-Luminous infrared galaxies (ULIRGs), and Wolf-Rayet galaxies. “Irregular galaxies often exhibit knots of starburst, often spread throughout the irregular galaxy... To ignite a starburst, it is necessary to concentrate a large amount of cool molecular gas in a small volume. Such concentrations and perturbations are strongly suspected to cause global starburst phenomena in major galaxy mergers, although the exact mechanisms are not fully understood.”⁵³⁵

Mergers in Elliptical and cD Galaxies

cD elliptical galaxies are probably the product of mergers. Most have multiple nuclei, and cD's are often located at the bottom of a cluster's potential well. Normal ellipticals also show evidence of mergers, especially boxy E's.

A majority of ellipticals (such as NGC 3923) have **shells** of stars partially enclosing them, which may be due to captured stars (from small faint galaxies) sloshing back in forth in their potential wells (*IMA2 p. 1014*). “Many and perhaps all large elliptical galaxies are strongly influenced by mergers.” (*IMA2 p. 1015*) The $r^{1/4}$ profiles are normal outcomes of mergers. However, dwarf ellipticals and dwarf spheroidals do not result from mergers.

⁵³⁴ <http://hubblesite.org/newscenter/archive/releases/2006/14/image/a/>

⁵³⁵ http://en.wikipedia.org/wiki/Starburst_galaxy

Binary Supermassive Black Holes

Galaxies, each having a black hole, that merge can lead to binary black holes orbiting each other, for instance probably NGC 6240 which appears to have double nuclei orbiting at about 0.8 kpc.⁵³⁶ Such mergers may be a pathway to the growth of supermassive black holes (*IMA2* p. 1016).

The Formation of Galaxies

A limited summary follows.

Eggen, Lynden-Bell, and Sandage ELS Collapse Model of the Milky Way

An early attempt to model our Milky Way Galaxy formation, 1962, proposed a “top-down” approach that involves the differentiation of a single, immense proto-Galactic cloud. But this is contradicted by 1/2 of outer halo stars having retrograde orbits. Another problem is the disparity in apparent ages among globular clusters (and other problems are mentioned).

Galaxy formation theory is not mature, and remains incomplete. We need a stellar birthrate function, based on star formation rate SFR and the initial mass function IMF. ...

The G-dwarf problem

This is an unexplained paucity of F and G main sequence stars near the Sun with low Z values.

A Dissipative Collapse Model...

This involves the timescale of collapse, whether it is free-fall or dissipative...

The Hierarchical Merger Model

This is a more recent bottom-up approach that improves on the ability to explain age and metallicity differences among the globular clusters. The proposal, fed by rising recognition of the importance of mergers, is that ELS top-down formation takes place, but so does bottom-up via a hierarchy of mergers. In the latter, disruption of globular clusters etc. leads to halo field stars... This would most strongly affect the center of the Galaxy, building the bulge more rapidly than the halo. Perhaps 90% of original clusters have been disrupted, especially the lower-mass ones and the high mass ones... The dSph galaxies may be proto-Galactic remnants.

Formation of the Thick and Thin Disks

An attempt is made to deduce the gravitational potential energy as a function of z-displacement and directed perpendicular to the disk, and thereby the local z-directed gravitational acceleration. This leads to the vertical scale height of the thick disk.

Thin disk forms after the thick disk. A feedback loop operates: If the thin disk becomes too thin, the mass density increases, supernovae increase, and the disk is reheated causing expansion. Most ongoing star formation occurs in a young, inner portion of the thin disk with scale height of < 100 pc.

⁵³⁶ <http://arxiv.org/abs/astro-ph/0106368>

Young Stars in the Central Bulge

These arise from recent mergers with gas-rich satellite galaxies—tidal interactions strip the gas, which settles into the disk, forming new stars...

Metallicity Gradients...

Hierarchical mergers predicts metallicity gradients in galaxies that have undergone a dissipative collapse. Such galaxies should exhibit a color gradient, namely that stars further from the center should be redder. However, this patterns could be disrupted by mergers etc., and in fact starburst galaxies have inverted color gradients and appear bluer in their centers.

Formation of Elliptical Galaxies

Many ellipticals have probably formed their stars early in the galaxy-formation process before a disk could form. Other ellipticals have formed from mergers of spirals that would also disrupt disks and relax to the elliptical pattern with an $r^{1/4}$ profile. Not fully explained is the distribution of globular clusters. Mechanisms are being evaluated with N-body simulations. A competing hypothesis is that ellipticals develop in the bottoms of the potential wells, even without interactions...

Galaxy Formation in the Early Universe

Early galaxy formation can be witnessed in distant galaxies. At 1 Mpc, the light we see was emitted 3 million years ago.

According to the 1978 **Butcher-Oemler effect**⁵³⁷, galaxies used to be bluer than they are today (suggesting more active new star formation back then). For more distant and therefore views of older times, there are fewer elliptical and lenticular galaxies relative to spirals, consistent with their later formation by mergers.

The HST 2004 Ultra Deep Field image was of only 3 arcmin square in Fornax, and includes some galaxies from only 400 to 800 Myr after the Big Bang. These distant galaxies exhibit an abundance of strange shapes and blue galaxies that may be the building blocks for the more regularly shaped galaxies we see closer to home.

The clumping of dark matter forming potential wells was likely critical to the formation of clusters of galaxies (IMA2 p. 1030) and will be discussed in Chap. 30.

⁵³⁷ http://en.wikipedia.org/wiki/Butcher-Oemler_Effect

The Structure of the Universe (Chapter 27)

Summary of Distance Measuring Methods and The Extragalactic Distance Scale

The two dimensions of an object's location on the celestial sphere (or "surface") are easy to measure, but the astrometric measurement of the third dimension, distance, is much more difficult. There are a number of methods that estimate distance to objects in astronomy, and they differ in applicable scale. The sequence of methods that apply at differing but overlapping distances establishes a **cosmological distance ladder** which is used to estimate distances to the farthest objects known.

Radar Ranging

On the closest scale, direct measurement of distance may be done precisely with **radar ranging**. This technique however is confined to objects in the solar system, as the transmitted plus reflected signal exhibit a combined fall-off $\propto 1/d^4$. Objects ranged have included Mercury, Venus, Moon, Mars, Jupiter and the Galilean satellites, Saturn including its rings and Titan, and a number of comets and asteroids.⁵³⁸ Saturn is at about 9 AU or about 1.3×10^{-4} LY, so that we may take 10^{-4} LY as the maximal distance measurements of this method.

Geometric or Trigonometric Parallax

This has been previously discussed (see Stellar Parallax). It is one of the gold standards outside the solar system, especially as performed by the highly precise Hipparcos Space Astrometry Mission ("High precision parallax collecting satellite", 1989–1993).⁵³⁹ Even better is expected from the upcoming and previously mentioned GAIA mission, slated to launch in 2013.⁵⁴⁰ Current maximal distance measurements with geometric parallax by Hipparcos has been 1 milliarcsec, equivalent to 1000 pc or 3000 ly, but more commonly ground measured distances to about 100 ly (including the Hyades star cluster⁵⁴¹ at 153 ly) are determined by this method.⁵⁴² It is said that GAIA will measure down to 10 or 20 microarcsec, the latter presenting $50 \text{ kpc} = 163,000 \text{ ly}$ (*IMA2* p. 59 and *here*⁵⁴³).

Main-Sequence Fitting ("Spectroscopic parallax")

This has been previously mentioned. It applies primarily to clusters of stars, for which a valid sample of stars are available at about the same distance. By comparing their H-R diagram to the H-R diagram of the absolute magnitude of well-established main-sequence stars (for example the Hyades cluster, and correcting presumably for any small amounts of redshift present), a displacement in magnitude may be observed between the two curves. The difference between the line of apparent magnitude of the cluster and the line of absolute magnitude of the main sequence line is the distance modulus used in the main sequence fitting. (Problems can arise from blue stragglers, older clusters that have few main-sequence stars, and juxtaposed stars that are not gravitationally bound in the cluster.) *IMA2* p. 1038 says this method extends to about 7 kpc or 2,000 ly, and ZXI suggests this technique is useful to about 10^5 ly. The open star cluster M13

⁵³⁸ http://en.wikipedia.org/wiki/Radar_astronomy

⁵³⁹ <http://en.wikipedia.org/wiki/Hipparcos>

⁵⁴⁰ <http://sci.esa.int/science-e/www/object/index.cfm?fobjectid=28820>

⁵⁴¹ http://en.wikipedia.org/wiki/Hyades_%28star_cluster%29

⁵⁴² <http://outreach.atnf.csiro.au/education/senior/astrophysics/parallaxlimits.html>

⁵⁴³ http://en.wikipedia.org/wiki/Gaia_%28spacecraft%29

(NGC 6205) is estimated by ZXI⁵⁴⁴ to be at 32,600 ly by this technique, whereas a recently published value is 6.8 kpc = 22,200 ly.⁵⁴⁵

Wilson-Bappu Effect

Another method involving analysis of stellar spectra is takes advantage of the Wilson-Bappu effect, in which widening of spectral lines correlates with absolute magnitude. For example, the Ca II K *absorption* line shows a well-defined centrally superimposed *emission* core with a quite narrow width that varies with absolute magnitude. The technique is useful over a range of 15 mag (IMA2 p. 1039). It is calibrated with nearby stars. Unfortunately, the dispersion of values is about 0.5 mag, rendering it too imprecise to significantly improve the cosmic distance ladder.⁵⁴⁶

Cepheid (and RR Lyrae) Variable Stars

Cepheids have been extensively discussed, in particular the Period-Luminosity-Color (PLC) relation. These supergiant stars are one of the gold standards for galaxies and clusters containing these stars (the Cepheids must be individually identifiable). The technique is also called **secular parallax** (a misnomer). Cepheids have been used to measure distance to M31, M100 (55 Mly), and other relatively nearby galaxies. Extinction problems are best addressed by observing in infrared wavelengths.

The much smaller older and fainter metal poor Type II cepheids must be distinguished from the supergiant metal-rich Classical Type I Cepheids.

The most distant known Cepheids (IMA2 p. 1041) are in NGC 3370⁵⁴⁷ at 29 Mpc or about 10^8 ly. ZXI says the method is useful to about 10^7 ly. Cepheids are an important method that spans the range of about 100 to 10^7 ly. The uncertainty is ± 0.4 mag (IMA2 p. 1051) or, from another source, about 7% (for nearby stars) to up to 15% (for distant stars).⁵⁴⁸

RR Lyrae variable stars also may be used to estimate distance because their luminosities are all about the same. They are more common than Cepheids, are metal poor Population II horizontal branch stars found in globular clusters, but also much less luminous than Cepheids..

⁵⁴⁴ <http://www.astro.washington.edu/users/ivezic/REU08/GDP/distance/background4.html>

⁵⁴⁵ Paust NEQ et al, "The ACS Survey Of Galactic Globular Clusters. VIII. Effects of Environment on Globular Cluster Global Mass Functions", *ApJ* Volume 139 Number 2 , 2010

⁵⁴⁶ http://en.wikipedia.org/wiki/Wilson%20Bappu_effect

⁵⁴⁷ <http://heritage.stsci.edu/2003/24/supplemental.html>

⁵⁴⁸ http://en.wikipedia.org/wiki/Cosmic_distance_ladder#Classical_Cepheids

Supernovae Type Ia

Core-collapse supernovae (SNe Types Ib, Ic, and II) have been used to estimate luminosity and therefore distance, but the uncertainties are high—about 25% for determining distance to the Virgo cluster (*IMA2 p. 1042*)—and Type II SNe at least are dimmer than Type Ia's.

However, the light curves of Type Ia SNe can be much more tightly fit to a calibration curve. These SNe are the best distance-determining standard candles for distances to greater than 1000 Mpc (3 billion ly). These highly uniform SNe occur when a binary white dwarf star, accreting matter from its companion red dwarf star, reaches the Chandrasekhar Limit of $1.4 M_{\odot}$. Once this critical mass is reached, the star undergoes a runaway nuclear fusion reaction and explodes. Because all Type Ia supernovae explode at about the same mass, their absolute magnitudes are all the same.⁵⁴⁹ The absolute magnitude at peak of Type Ia SN has been found to be $\langle M_B \rangle \cong \langle M_V \rangle \cong -19.3 \pm 0.3$.

In the upper graph to the right,⁵⁵⁰ the absolute blue magnitude M_B is plotted over 80 days (in the star's rest frame). Most of the stars fall in the narrow yellow band (not all stars are plotted), and there are few outliers. The lower graph is the same light curves rescaled. Perlmutter 2003 states “The nesting of the light curves suggests that one can deduce the intrinsic brightness of an outlier from its time scale. The brightest supernovae wax and wane more slowly than the faintest... Simply by **stretching** the time scales of individual light curves to fit the norm, and then scaling the brightness by an amount determined by the required time stretch, one gets all the type Ia light curves to match.”

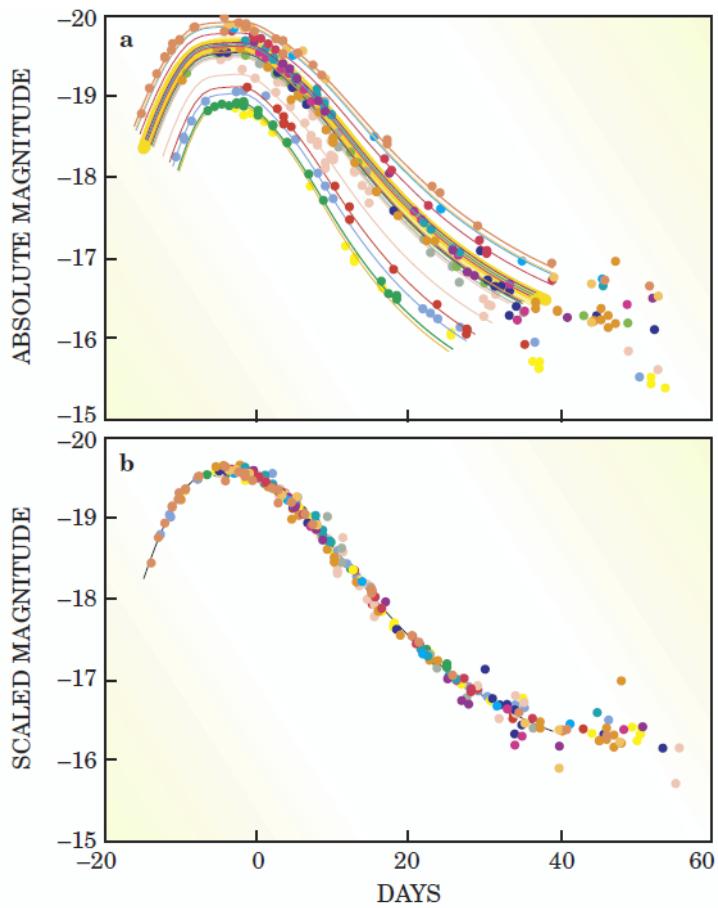
Catching the SN at peak luminosity is desirable but not apparently essential for inferring the peak luminosity. This inference involves either the aforementioned “stretch” method, or the **multicolor light curve shape method (MLCS)**, in which the shape of the light curve (taken after the initial explosion) is compared to a family of parameterized template curves that will allow determination of the absolute magnitude at peak brightness. The latter has the advantage of allowing for dust-related extinction and reddening of transmitted light.

Once the scaled absolute blue magnitude and the apparent blue peak magnitude is determined, the distance is calculated as usual with

$$d = 100^{(m - M - A + 5)/5} \text{ pc}$$

Type Ia SNe are about 13.3 magnitudes brighter than Cepheid variable, thus the technique is capable of reaching out more than 500 times further. The uncertainty in distance modulus is about ± 0.1 mag, and the uncertainty overall in distance estimation by this method is as low as 5%. The range is > 1000 Mpc. (*IMA2 p. 1044*)

However, Type IA supernovae (SNe) are quite rare. The last observed SN of any type in the Milky Way was the 1604 “Kepler’s Supernova” (SN 1604, type uncertain but apparently a Type IA).⁵⁵¹ Before Kepler’s SN, another Type IA SN, SN 1572 (“Tycho’s Supernova”), occurred in 1572.⁵⁵²



⁵⁴⁹ http://en.wikipedia.org/wiki/Cosmic_distance_ladder#Type_Ia_light_curves

⁵⁵⁰ Perlmutter, Saul, “Supernovae, Dark Energy, and the Accelerating Universe”, Physics Today, 56, April, 53 (2003).

⁵⁵¹ <http://chandra.harvard.edu/photo/2007/kepler/index.html>

⁵⁵² http://en.wikipedia.org/wiki/SN_1572

One volumetric survey estimate,⁵⁵³ based on the SN Ia **rate** found from the Sloan Digital Sky Survey II (SDSS-II) Supernova Survey (which found 516 SN Ia's at redshift $z < 0.3$), is that the SN rate at a mean redshift of $z = 0.12$ is

$$r_V = 2.69 \times 10^{-5} \text{ SN yr}^{-1} \text{ Mpc}^{-3} (H_0/(70 \text{ km s}^{-1} \text{ Mpc}^{-1}))^3$$

Novae

These are more common and may also be used like SNe to estimate distance (*IMA2* p. 1044), typically using the time from peak M_V to 2 magnitudes less. This method extends to 20 Mpc, just past the distance to the center of the Virgo cluster and similar to the maximal distance for Cepheids. The uncertainty is ± 0.4 mag. Of course, the range is much smaller than Type Ia SNe.

Brightest Giant H II Regions; Brightest Red Supergiants

These can be used to provide a secondary standard candle (i.e., one requiring a galaxy at known distance for calibration. Angular size vs. galaxy apparent magnitude are measured for the object and compared to the same quantities in galaxies with known distance, etc. This method is not very accurate.

Similarly, the 3 brightest individually resolved Red Supergiants in a galaxy seem to have about the same absolute $M_V = -8.0$, and this can be the basis of distance estimate. This method is limited to about 7 Mpc.

⁵⁵³ Dilday B, et al, "Measurements of the Rate of Type Ia Supernovae at Redshift $z < \sim 0.3$ from the SDSS-II Supernova Survey", *ApJ* 713:1026-1036, 2010, <http://arxiv.org/abs/1001.4995> Error range omitted.

Globular Cluster Luminosity Function GCLF

This statistical method is more accurate than measuring the top 3 red supergiants, and works best for giant ellipticals with many globular clusters. For a globular star cluster, the uncertainty is about 0.4 mag, and the method reaches out to 50 Mpc. However, uncertainties remain because there is no universal globular cluster luminosity. The graph that follows⁵⁵⁴ demonstrates the peak number of GCs found over all luminosities for GCs around 4 elliptical galaxies in the Virgo Cluster, as well as the standard deviation in the peak.

The “turnover” absolute magnitude has been found with other samples to be $M_0 \cong -6.6$.

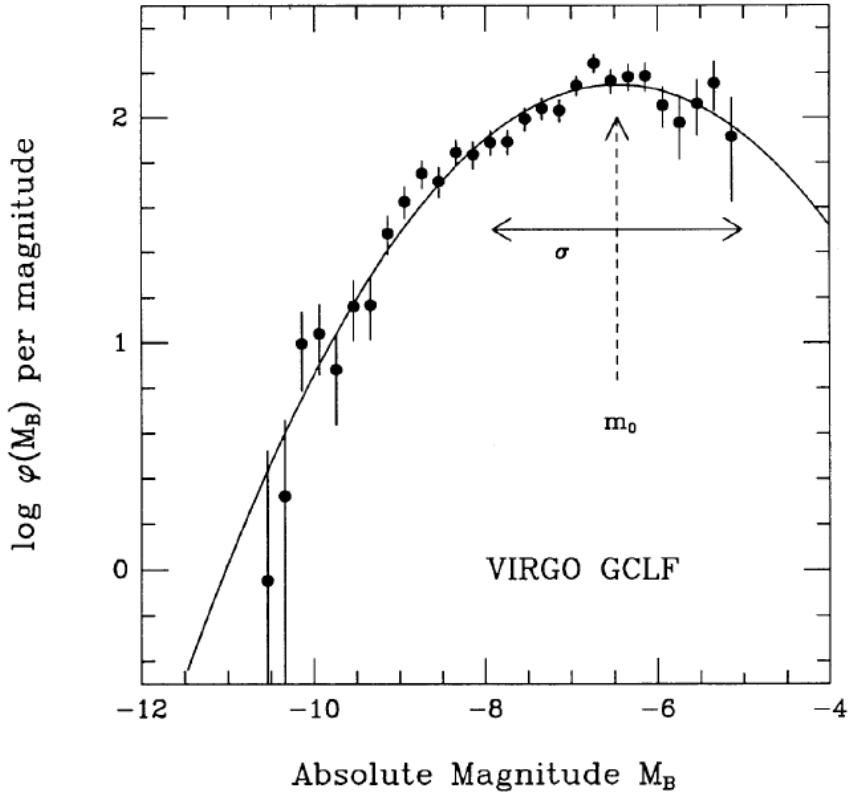


FIG. 7—The luminosity function for the globular clusters in four Virgo ellipticals combined (NGC 4365, 4472, 4486, 4649), from Harris et al. (1991). This composite GCLF was constructed from a total sample of about 2000 clusters brighter than $B=26.2$ (the vertical scale units shown are arbitrary). Here ϕ is the relative number of clusters per unit magnitude interval, plotted against absolute magnitude M_B assuming $d(\text{Virgo}) \simeq 17$ Mpc. A Gaussian interpolation curve is superimposed, with a “turnover” or peak indicated by the dashed arrow (m_0), and a standard deviation $\sigma=1.45$ mag.

This technique has an uncertainty is ± 0.4 mag (IMA2 p. 1051).

⁵⁵⁴ Jacoby GH, “A Critical Review of Selected Techniques For Measuring Extragalactic Distances”, Publications of the Astronomical Society of the Pacific, 104: 599-662, 1992 August,
<http://ned.ipac.caltech.edu/level5/Jacoby/frames.html>

Planetary Nebula Luminosity Function PNLF

A similar statistical approach can be used with a collection of a galaxy's planetary nebulae. The value $M_{5007} = -4.53$ expresses the cutoff (maximum) absolute magnitude at the 500.7 nm [O III] forbidden line found in all Population II PNs, a value which may be used as the standard candle. This is reliable in finding distances to elliptical galaxies up to 20 Mpc (*IMA2* p. 1046, see also *here*⁵⁵⁵) or up to 50 Mpc (*IMA2* p. 1051). The uncertainty is ± 0.3 mag (*IMA2* p. 1051)

Surface Brightness Fluctuation Method

This technique looks at a galaxy as a whole, and can be used up to 100 Mpc, 125 Mpc with HST. It is best done with a CCD detector. The amount of rms pixel to pixel variation in brightness is inversely proportional to distance. (*IMA2* p. 1047 and *here*⁵⁵⁶) This technique has an uncertainty is ± 0.3 mag (*IMA2* p. 1051).

Tully-Fisher Relation (Spirals $V_{\max} - L$)

Discussed previously, applies to spiral galaxies, relates spiral galaxy luminosity to maximum rotation velocity. It is accurate (± 0.4 mag in IR). Despite its great range, > 100 Mpc, nearby spirals can also be calibrated with Cepheids. This has helped to map the 3-D dimensions of the Virgo and Coma Clusters. It is widely used in measuring extragalactic distances. (*IMA2* p. 1048 and *here*⁵⁵⁷) This technique has an uncertainty is ± 0.4 mag (*IMA2* p. 1051).

D versus σ Relation ($D - \sigma$)

This applies to elliptical galaxies. As previously mentioned, the Faber-Jackson relation $L_B \propto \sigma_0^\alpha$ correlating luminosity and velocity dispersion σ_0 is subject to excessive scatter for elliptical galaxies and not useful for estimating distance. However, the relation between galaxy diameter D and velocity dispersion σ for a galaxy's stars is tighter. D here is the angular diameter out to a surface brightness (isophote) of 20.75 B-mag arcsec⁻². D is inversely proportional to the galaxy's distance. This method then give a standard ruler rather than a standard candle, in that a measured σ_0 predicts a certain diameter D, by $\log_{10}D = 1.333 \log_{10}\sigma + C$. Here, C depends of the distance to the galaxy cluster. This is also useful for determining relative distances between two galaxy clusters. It can be used to show that Coma Cluster is more than 5 times as far as the Virgo Cluster. However, this method has not yet been calibrated with Cepheids. (*IMA2* p. 1048 and *here*⁵⁵⁸) This technique has an uncertainty is ± 0.5 mag (*IMA2* p. 1051) and a range of > 100 Mpc.

The term “**fundamental plane**” is sometimes used for this technique as applied to normal elliptical galaxies. The plane in question is not a normal spatial plane, but a 2-dimensional abstract plane within a 3-dimensional abstract “space” having coordinates of effective radius, average surface brightness, and central velocity dispersion. “Any one of the three parameters may be estimated from the other two, as together they describe a plane that falls within their more general three-dimensional space.”⁵⁵⁹ The Luminosity may be given by an empirical formula expressing it as a function of dispersion and effective radius. Alternatively, the effective radius may be given by an empirical formula expressing it as a function of dispersion and surface brightness (*IMA2* p. 988)

⁵⁵⁵ Planetary nebula luminosity function:

- http://en.wikipedia.org/wiki/Planetary_nebula_luminosity_function
- see also <http://arxiv.org/abs/0708.4292>

⁵⁵⁶ http://ned.ipac.caltech.edu/level5/Jacoby/Jacoby9_1.html

⁵⁵⁷ Tully-Fisher Relation:

- <http://www.noao.edu/staff/shoko/tf.html>
- http://en.wikipedia.org/wiki/Tully%20%80%93Fisher_relation

⁵⁵⁸ http://en.wikipedia.org/wiki/Cosmic_distance_ladder#D-.CF.83_relation

⁵⁵⁹ http://en.wikipedia.org/wiki/Fundamental_plane_%28elliptical_galaxies%29

Brightest Galaxies in Clusters (BCG)

The brightest galaxy in a cluster of galaxies can be used to obtain the cluster's distance. The brightest value has been found to average absolute visual magnitude $M_V = -22.83 \pm 0.61$ mag, approximately 3 magnitudes brighter than the peak brightness of Type Ia SNe. This method should be good to over 4000 Mpc, and thus allow viewing of light emitted >13 billion years ago. However, galaxies may have evolved in intrinsic brightness over cosmological time periods, and galaxies at great distances may be different, for instance by mergers, or they may be smaller,⁵⁶⁰ so caution is needed in applying this method.

The Expansion of the Universe

Vesto M Slipher (1875–1969) first discovered in 1914 that most nebulae were redshifted and therefore moving away from us. (although the relatively nearby Andromeda M31 galaxy is blue shifted). He also found that most nebulae are moving apart from each other, indicating a general **expansion of the universe**.

Hubble Law

Hubble discovered Cepheids in M31 in 1925. His publication in 1929 included the Hubble Law or Relation: the velocity of distant galaxy recession is proportional to the distance. This relation is now expressed in non-relativistic form as $v=H_0 d$. The distance d is given in Mpc and recession velocity v in km s^{-1} , so H_0 has units of $\text{km s}^{-1} \text{Mpc}^{-1}$. He and assistant Milton Humason compiled many distances and redshifts, working for instance with the H and K lines of Ca II.

The factor H_0 applies only to the present time and is commonly expressed as $100h \text{ km s}^{-1} \text{ Mpc}^{-1}$, where h is a dimensionless multiplier derived from various sources. (It is of course not the same as the Planck constant h). The WMAP value of h , namely h_{WMAP} , is currently accepted as one the best estimates:

$$h = h_{\text{WMAP}} = 0.710^{+0.04}_{-0.03}$$

Therefore, when $z \ll 1$, $d \cong cz/100h \text{ Mpc}$, or $d \cong cz/71 \text{ Mpc}$. This is the "**naïve**" **Hubble distance**, applicable only for small z .

Recessional Velocity

According to *IMA2* p. 1054, a galaxy's **recessional velocity** is not due to its movement through space but instead to an expansion of space itself which carries the galaxy along. Thus it differs from **peculiar velocity**. This motion of the galaxies due to expansion is called the **Hubble Flow**. Note that gravitationally bound systems do not expand with the Hubble Flow (*IMA2* p. 1056).

Cosmological Redshift

The redshift attributable to expansion of the universe is called the **cosmological redshift**, and is due to the corresponding expansion of the wavelength of the light emitted—thus it is not accurate per General Relativity to use the Doppler shift equations developed for a static Euclidean universe. However, astronomer s use the relativistic Doppler formula for convenience to translate a measured redshift parameter z into a peculiar velocity that would apply if the redshift were due to relative motion rather than to the expansion of spacetime.

Cosmological redshift attributable to Hubble expansion is so named to distinguish it from more familiar Doppler redshift (due to peculiar velocities of galaxies, etc.) and from gravitational redshift (arising when light climbs out of a deep potential well).

The following is an abbreviated discussion.

Recall that redshift parameter z is defined:

⁵⁶⁰ http://www.nasa.gov/mission_pages/hubble/science/far-protocluster.html
Page 215 of 249

$$z = (\lambda_{\text{obs}} - \lambda_{\text{rest}}) / \lambda_{\text{rest}} = \Delta\lambda / \lambda_{\text{rest}} = \lambda_{\text{obs}} / \lambda_{\text{rest}} - 1 = v_r/c$$

or

$$z = (f_{\text{rest}} - f_{\text{obs}}) / f_{\text{obs}} = \Delta f / f_{\text{obs}} = f_{\text{rest}} / f_{\text{obs}} - 1 = v_r/c$$

The relativistic Doppler shift is given by

$$v_{\text{obs}} = v_{\text{rest}} \frac{\sqrt{1 - u^2/c^2}}{1 + v_r/c} - 1$$

where v_r is the observed radial velocity, $v_r = u \cos \theta$, and θ = angle between line to observer and the direction of movement at velocity u .

The redshift for purely radial motion is given as

$$z = \sqrt{\frac{1 + v_r/c}{1 - v_r/c}} - 1$$

and distance is given approximately by

$$d \cong \frac{c}{H_0} \frac{(z+1)^2 - 1}{(z+1)^2 + 1}$$

When $z \ll 1$, $d = cz/H_0$.

An excellent general-relativity-aware JavaScript web-page calculator by EL Wright is found [here](#).⁵⁶¹ One inserts z (plus optionally values for H_0 , Ω_M , Ω_{vac} , and whether) and computes comoving radial distance, time since the Big Bang, and light travel (elapsed) time.

The following graphs show various distances in Gly (or age in Ga for lookback time) as a function of cosmological redshift z .⁵⁶² (Graph curves in the left diagram are named from top to bottom in the legend at lower right of this graph.) Note that the Naïve Hubble graph (dot-dash line) is always a straight line, because for this graph $D_{\text{naive}} = cz/H_0$. Note log-log scales for second graph axes. Here,

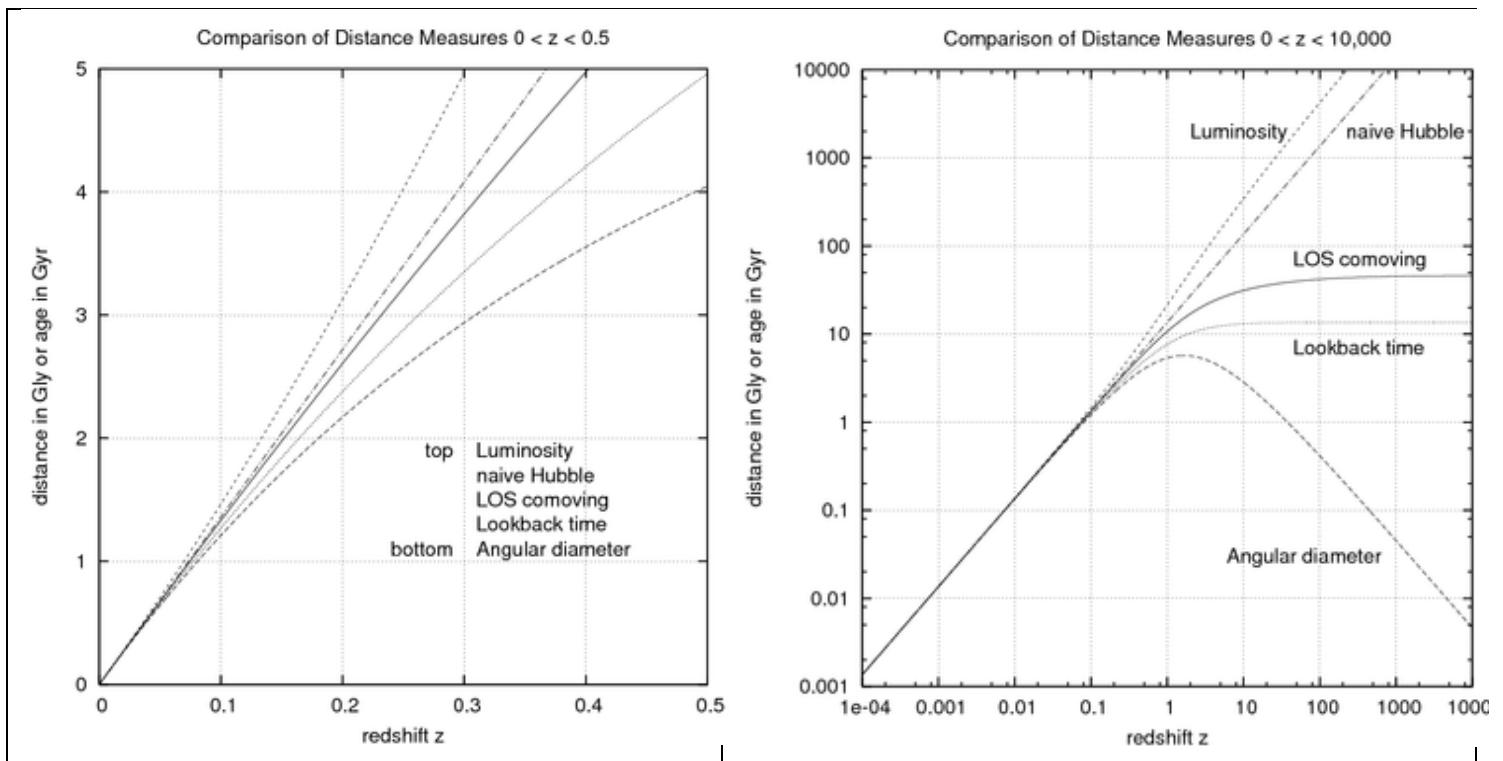
- Naïve Hubble $D_{\text{naive}} = cz/H_0$.
- LOS comoving D = line-of-sight comoving radial distance=proper motion distance;
- Luminosity DL based on luminosity vs. flux for known object, where Flux = Luminosity/(4*pi*DL²) where $DL = DA^*(1+z)^2$
- Look back time t_L =light travel distance = $c(t_0-t_{\text{emit}})$ (hard to determine because t_{emit} uncertain).
- Angular diameter DA= angular size distance = x/θ (x =actual size, θ = angular size), $DL = DA^*(1+z)^2$

⁵⁶¹ <http://www.astro.ucla.edu/~wright/CosmoCalc.html>, which is described in Wright Edward L., "A Cosmology Calculator for the World Wide Web", *The Publications of the Astronomical Society of the Pacific*, Volume 118, Issue 850, pp. 1711-1715, 12/2006. The article spells out how the general relativistic calculations are made. I have not studied general relativity and have not taken the time to review the details of these calculations.

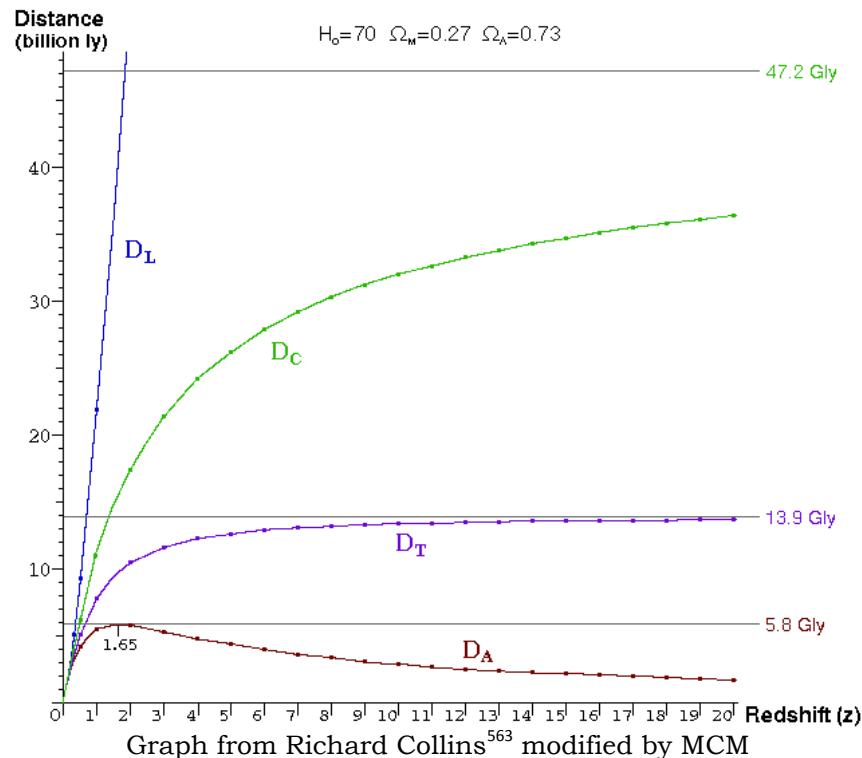
⁵⁶² http://en.wikipedia.org/wiki/Distance_measures_%28cosmology%29

See webpage for detailed explanation, not all of which I understand. See also and

http://www.astro.ucla.edu/~wright/cosmo_02.htm and <http://www.atlasoftheuniverse.com/redshift.html>



Another graph follows, showing the full range of cosmological redshifts for objects currently visible with HST (about $z = 15$, neglecting the CMB at $z \approx 1100$):



Note linear scales for axes.

D_L = Luminosity Distance, D_C = Comoving Distance,

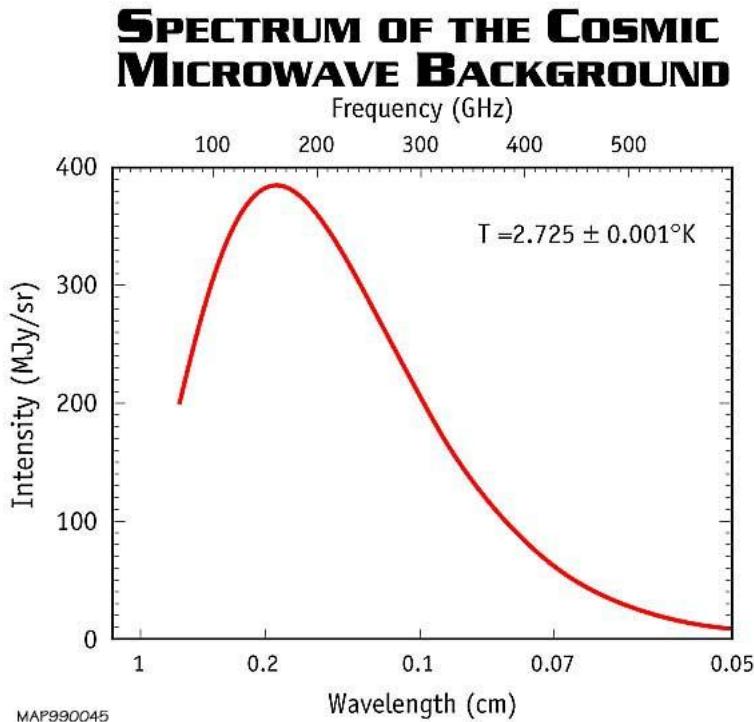
D_T = Light Travel Time Distance (Lookback Time distance), and D_A = Angular Diameter Distance

⁵⁶³ <http://www.atlasoftheuniverse.com/redshift.html>

"The Comoving Distance [Dc] is the distance scale that expands with the universe. It tells us where the galaxies are now even though our view of the distant universe is when it was much younger and smaller [and the object much closer]. On this scale the very edge of the visible universe is now about 47 billion light years from us although the most distant galaxies visible in the Hubble Space Telescope will now be about 32 billion light years from us."⁵⁶⁴

Big Bang and Cosmic Microwave Background CMB

The Big Bang is the inferred past beginning of an explosion taking place where all matter and space was consolidated into a tiny size, perhaps a mere point. The Big Bang left a remnant in the form of blackbody radiation of this intensely hot explosion. Due to the expansion of space as the light from this event has propagated, the wavelengths of this intensely energetic light emitted have become elongated, causing an effective cooling of the radiation, so that it now has a blackbody $T_{\text{eff}} = 2.725 \text{ K}$ (graph from NASA⁵⁶⁵ follows).



If we assume that the expansion has proceeded at uniform velocity (actually an invalid assumption) we get a crude estimate of the **Hubble Age** or **Hubble Time** t_H , the time it takes to play the Big Bang explosion in reverse back to its origin. This time $t_H = 13.8 \text{ Gyr}$. (*IMA2 p. 1058*)

Groups and Clusters of Galaxies

The universe is still not of uniform structure at a scale somewhat greater than that of galaxies. Instead, galaxies often occur gravitationally bound and orbiting the barycenter of galaxy groups and clusters. The structure of the universe consists of galaxies, groups of galaxies, clusters of galaxies, and clusters of clusters (Superclusters).

Groups of Galaxies: are the smallest bound aggregate of galaxies, with typically no more than 50 galaxies in a diameter of $1.4 h^{-1} \text{ Mpc}$ and an average mass of approximately $2 \times 10^{13} h^{-1} M_\odot$. The spread of velocities for the

⁵⁶⁴ ibid.

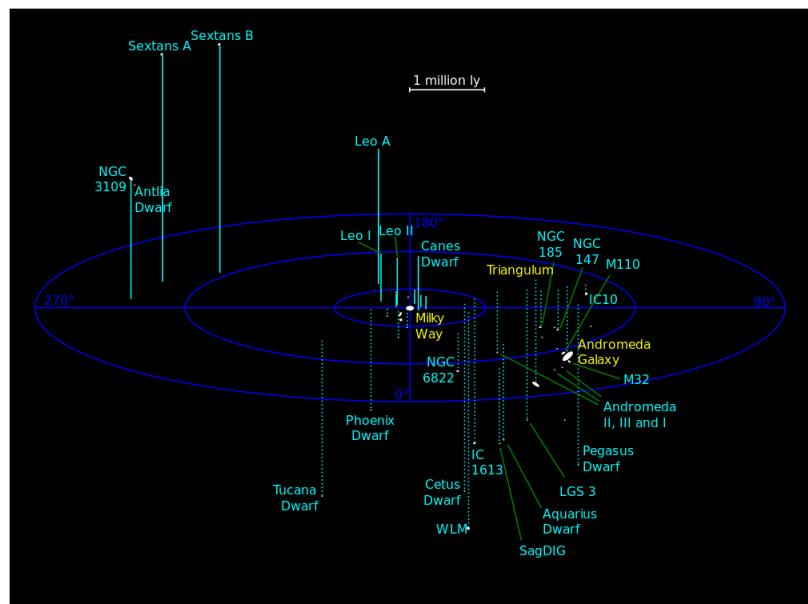
⁵⁶⁵ http://map.gsfc.nasa.gov/universe/bb_tests_cmb.html

individual galaxies is about 150 km/s. Larger and more massive galaxy systems are sometimes also classified as galaxy groups. The mass to light ratio is typically about $260h M_\odot/L_\odot$, indicating much dark matter.

The Milky Way is contained in the **Local Group** of more than 54 galaxies, see below. (IMA2 p. 1058 and [here](#)⁵⁶⁶)

The graph to the right depicts the relative 3D positions of the galaxies in the Local Group.⁵⁶⁷ The graph is centered on the MW, but the MW is not at the barycenter (it is 462 kpc from the MW). Note the wide separation between MW and M31 (Andromeda Galaxy).

Other groups are perhaps less well known, but include the IC 342/Maffei Group, the M81 Group, and the Centaurus A/M83 Group. (The two groups with “/” in the name each consist of 2 subgroups, thus the complex naming). The Sculptor group is the closest group at 1.8 Mpc, and the M81 group is at 3.1 Mpc. Several additional groups are within 10 Mpc of the Local Group.



Galaxy Clusters: are larger than groups, although there is no sharp dividing line between the two. Clusters hold from 50 (a **poor** cluster) to up to many thousands of galaxies (a **rich** cluster) within a span of about $6h^{-1}$ Mpc, and with a typical mass of $1 \times 10^{15} h^{-1} M_\odot$. Velocity dispersions are 800 to up to 1000 km s⁻¹. The mass to light ratio is typically about $400h M_\odot/L_\odot$, again implying much dark matter. Clusters can be spherical or irregular. (IMA2 p. 1058). Most galaxies are found in small groups or poor cluster (akin to the fact that there are more dwarfs than giant galaxies). (IMA2 p. 1061)

Notable nearby galaxy clusters include the Virgo cluster, Fornax Cluster, Hercules Cluster, and the Coma Cluster.⁵⁶⁸

Zero Velocity Surface

The edge of a cluster or group is not well defined visually or otherwise, and thus the concept of **zero velocity surface ZVS** is useful to define their spatial limits. (The term ZVS arises in the problem of N-body orbits.) It is especially useful with loosely bound orbiting bodies such as galaxies. The ZVS is the mathematically-defined 3-D surface that is the theoretical eventual limit of excursions of the galaxies in the cluster, given their current velocities and directions. In other words, it defines a surface boundary that the galaxies cannot cross. (IMA2 p. 1059 and [here](#)⁵⁶⁹, a complex topic that I have not studied)

The Local Group

This is more than 54 nearby galaxies, located within about 1 Mpc from the MW. The ZVS is at about 1.2 Mpc from the barycenter. The Local Group includes the MW, the LMC, the SMC, M31 (Andromeda), the Triangulum Galaxy, and an assortment of dwarf ellipticals and dwarf spheroidals, etc., some of which are in the Magellanic Stream. It is part of the Virgo Supercluster. The MW and M31 are the largest galaxies and

⁵⁶⁶ http://en.wikipedia.org/wiki/Galaxy_groups_and_clusters

⁵⁶⁷ http://en.wikipedia.org/wiki/Local_Group

⁵⁶⁸ http://en.wikipedia.org/wiki/Galaxy_cluster

⁵⁶⁹ <http://farside.ph.utexas.edu/teaching/336k/newton/node125.html>

widely separated, altogether providing 90% of the luminosity of the group. The Local Group is in a state of collapse, and in particular the MW and M31 will collide (but only after 6.3 Ga), perhaps forming a giant elliptical. It is inferred that after the Big Bang, the galaxies of the Local Group were initially in recession from each other, but gravitational attraction eventually stopped this expansion, and they began approaching each other. The total mass of MW and M31 combined is calculated by orbital considerations to be $4 \times 10^{12} M_{\odot}$. Their mass to light ratio is about $57 M_{\odot}/L_{\odot}$ but luminous matter (in the MW at least) only explains about 3 M_{\odot}/L_{\odot} , again implying much dark matter. (IMA2 p. 1060-1)

Virgo Cluster

This rich irregular cluster is in the constellations of Virgo and Coma Berenices. It covers 10 degrees square in the heavens. The center is about 16 Mpc from us. It comprises about 2250 small and large galaxies in a span of about 3–4 Mpc.⁵⁷⁰ The four brightest galaxies are all giant elliptical galaxies, 3 of which are close to the center of the cluster. These include M87 (Virgo A or NGC 4486), as well as M84 (NGC 4374) and M86 (NGC 4406).

M87 is a giant E1 elliptical active galaxy having a mass of about $6 \times 10^{12} M_{\odot}$ enclosed in 50 kpc radius, about 200 times that of the MW (IMA2 p. 1062 and *here*⁵⁷¹). For M87, the enclosed mass M_r increases linearly with radius out to about 300 kpc, a finding consistent with dark matter accounting for 99% of the mass in this galaxy. (IMA2 p. 1063)

Coma Cluster (Abell 1656)

This is the nearest rich and regular cluster, located in the constellation Coma Berenices. It has about 10,000 galaxies (most of which are dwarf ellipticals and S0s, with only 15% spirals and irregulars). It is about 90 Mpc away, about 5 times further than the Virgo Cluster. It is at a redshift $z=0.0231$, yielding a light travel time of about 0.3 Ga, and has an estimated mass = $7 \times 10^{14} M_{\odot}$. Its brightest members are the giant cD ellipticals NGC 4874 and NGC 4889, both near the center. Younger spirals are few and probably peripheral.⁵⁷² Fritz Zwicky first observed that there is not nearly enough luminous matter to account for observed orbits of the Coma galaxy cluster, and posited dark matter. The mass to light ratio can be simplistically calculated at about $660 M_{\odot}/L_{\odot}$ (but actually is about half that, IMA2 p. 1066)

Evolution of Galaxies

By analysis of distant and therefore older galaxy clusters , such as the rich cluster CL 0939+4713 (Abell 851) at $z=0.41$ (distance = $6.4 h^{-1}$ corresponding to a view from 9 Ga ago), we can infer how galaxies in clusters may have evolved. For this much older cluster, many spirals are seen in the center, whereas in the much more recent views of Coma Cluster, few central spirals are seen. (IMA2 p. 1064) It is theorized that the ellipticals form from tidal interactions and hierarchical protogalaxy mergers, leaving ellipticals and SOs.

An even more distant cluster , which includes 3C 324 and is in constellation Serpens, has $z=1.2$ and is therefore viewed at 9 Ga ago.⁵⁷³ It shows a much more disorderly arrangement of galactic fragments, perhaps representing future or former spirals. The HubbleSite description⁵⁷⁴ states: “Very few of the cluster's members are recognizable as normal spiral galaxies (like our Milky Way), although some elongated members might be edge-on disks. Among this zoo of odd galaxies are ‘tadpole-like’ objects, disturbed and apparently merging systems dubbed ‘train-wrecks’, and a multitude of faint, tiny shards and fragments, dwarf galaxies or possibly an unknown population of objects. However, the cluster also contains red galaxies that resemble mature examples of today's elliptical galaxies. Their red color comes from older stars that must have formed shortly after the Big Bang.”

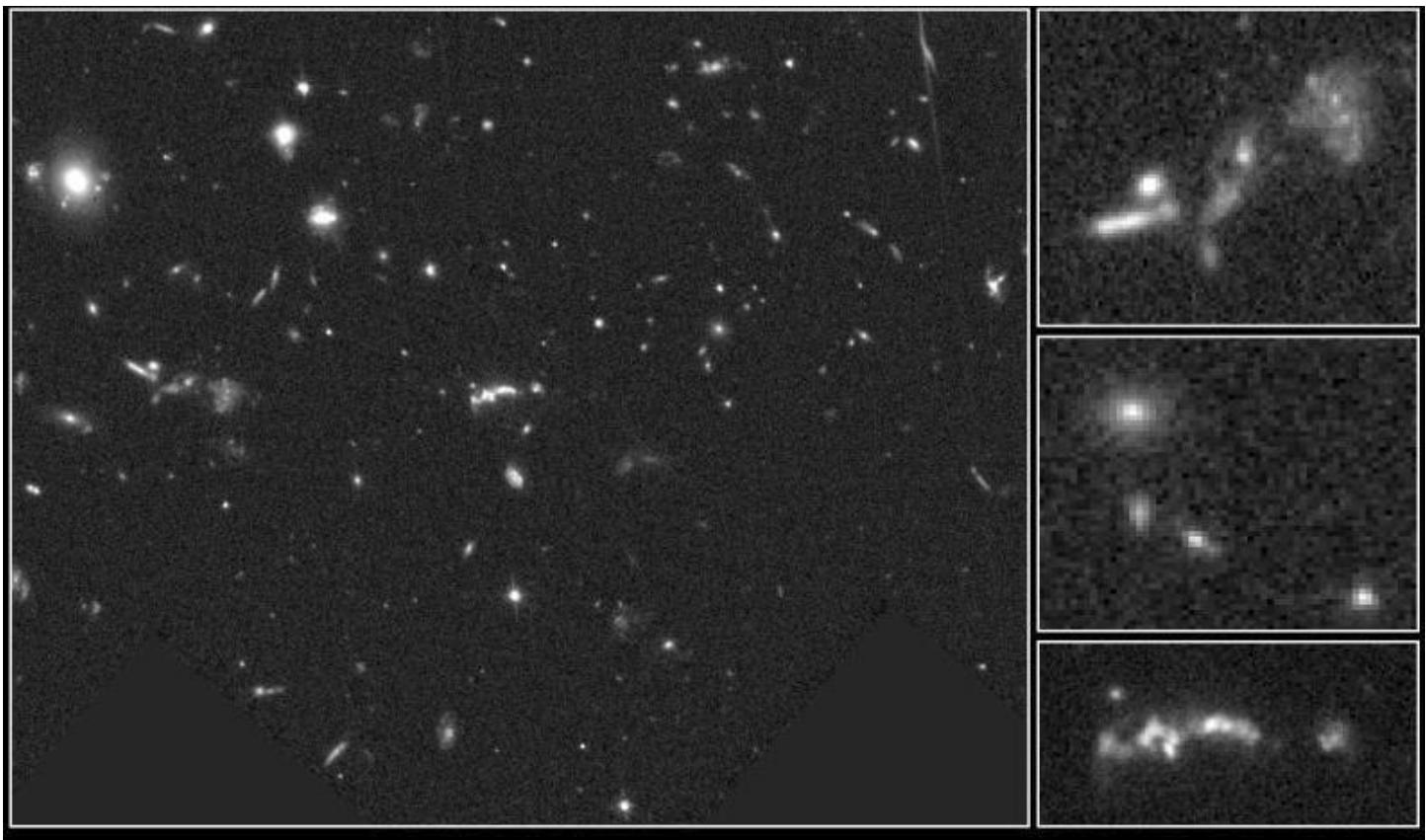
⁵⁷⁰ http://en.wikipedia.org/wiki/Virgo_cluster

⁵⁷¹ http://en.wikipedia.org/wiki/Messier_87

⁵⁷² http://en.wikipedia.org/wiki/Coma_Cluster

⁵⁷³ <http://hubblesite.org/newscenter/archive/releases/1994/52/>

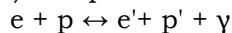
⁵⁷⁴ <http://hubblesite.org/newscenter/archive/releases/1994/52/image/a/>



Distant Galaxy Cluster Around 3C 324 HST • WFPC2

Hot Intergalactic Intracluster Gas and Intercluster Medium ICM

Some (but not nearly all) of the missing mass per Zwicky's calculation of mass to luminosity relationship is attributable to **intracluster gas** in the **intracluster medium ICM**. This is very hot, fully ionized, and emits in the X-ray wavelengths, and "filling the cluster gravitational potential wells". The X-ray luminosities for cluster ICM are 10^{36} to 10^{38} W. The Virgo and Coma Clusters have, resp., about 5×10^{13} and $3 \times 10^{13} M_{\odot}$ in ICM gas mass. The radiation is produced by the thermal bremsstrahlung process (free-free radiation), where the bremsstrahlung X-rays are created by the acceleration of free electrons in the electric fields of ions (mostly protons) in a plasma. The reaction is diagrammed as:



According to ZXI, there is little relatively optical emission because there are no optical transitions and the gas is optically thin (thus not like a black body).

For fully ionized H, the energy emitted by thermal bremsstrahlung per unit volume per unit time between frequencies ν and $\nu + d\nu$ is given by (IMA2 p. 1066):

$$l_v d\nu = 5.44 \times 10^{-52} (4\pi n_e^2) T^{-1/2} e^{-hv/kT} d\nu \text{ W m}^{-3}$$

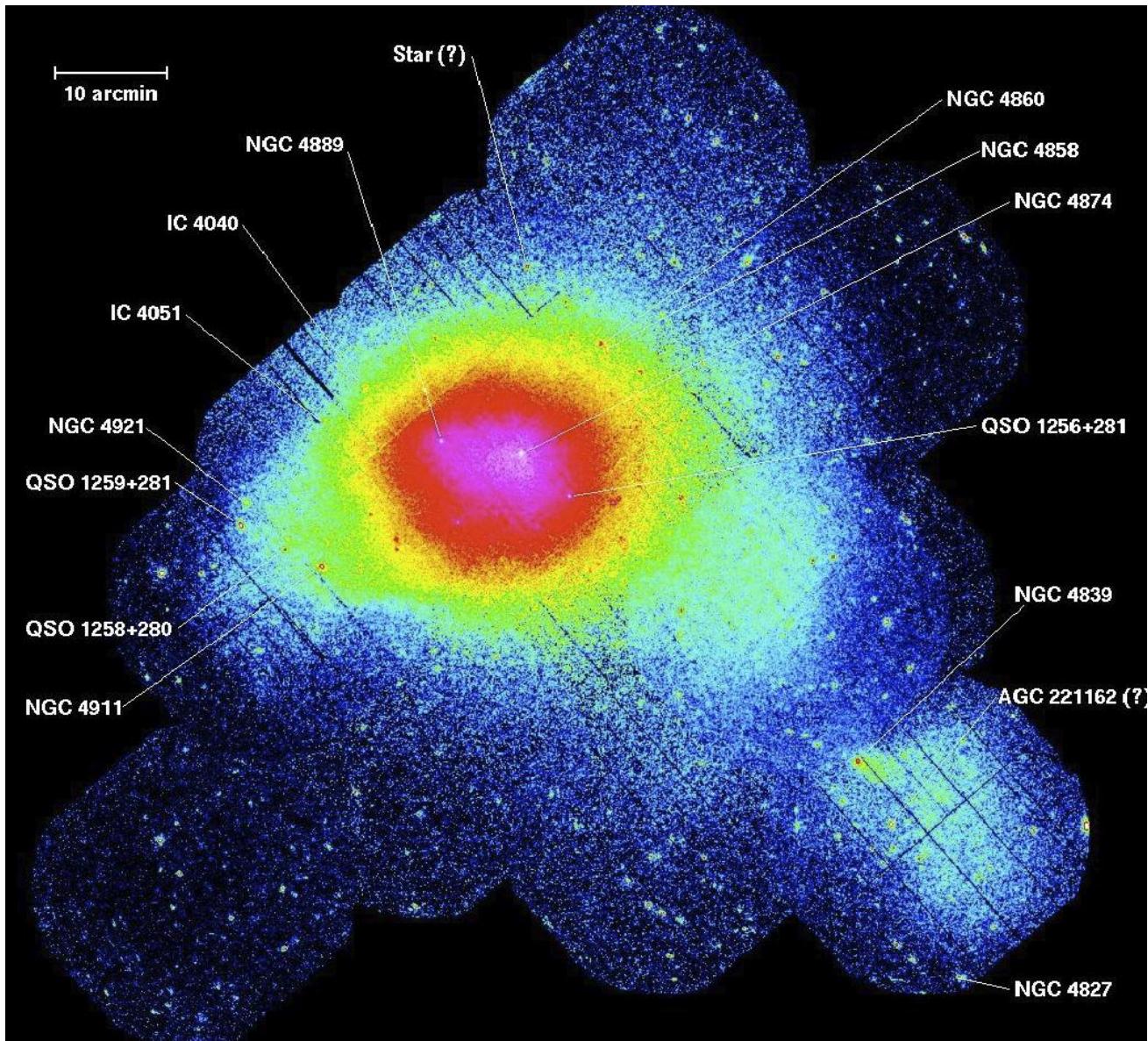
and for all wavelengths is

$$L_v = 1.42 \times 10^{-40} n_e^2 \text{ W m}^{-3}$$

An ESA/XMM image of the X-ray emission in the 0.3-2.0 keV energy band from the ICM of the Coma Cluster follows.⁵⁷⁵ The irregular distribution of the ICM may indicated mergers of several subclusters. The brightest

⁵⁷⁵ from the ESA/XMM-Newton mission, EPIC-pn camera, U. Briel, MPE Garching, Germany, color calibration not provided, http://xmm.esac.esa.int/external/xmm_science/gallery/public/level3.php?id=47, published in

region, shown in red and yellow, is centered on NGC 4874 and NGC 4889, giant cD elliptical galaxies, but clearly there is much emission outside the galaxies that are indicated.



For the Coma Cluster, as observed with NASA's HEAO 1 A-2 (High Energy Astronomy Observatories, Cosmic X-ray Experiment CXE) and OSO 8 (8th Orbiting Solar Observatory), the best fit for the observations as thermal bremsstrahlung process yields effective temperature $T = 7.6 \text{ keV} = 88 \times 10^6 \text{ K}$ (*here*⁵⁷⁶ and *IMA2 p. 1068*). The number of free electrons per m^3 is about 300, thus much less dense than giant molecular clouds (as previously described). The total mass in the ICM, estimated by the formulas just given and assuming 1 proton per electron, is $1 \times 10^{14} \text{ M}_\odot$, which overestimates compared to the preferred estimate of $3 \times 10^{13} \text{ M}_\odot$. Even the former is only a few percent of the total mass of the cluster.

The ICM also exhibits emission lines of highly ionized Fe, Si, and Ne, indicating that it has been enriched by stellar nucleosynthesis, perhaps through dispersion from high dynamical friction during mergers, or by bursts of star formation.

Briel UG, et al, "A Mosaic of the Coma Cluster of Galaxies with XMM-Newton", *Astronomy and Astrophysics*, v.365, p.L60-L66 (2001)

⁵⁷⁶ Henriksen MJ and Mushotzky RF, "The X-ray spectrum of the Coma Cluster of galaxies", *Astrophys. J.* 302, 287, 1986

The Existence of Superclusters, Bubbles, Filaments, and Voids

Superclusters exist and are up to 100 Mpc in size. Our Local Supercluster contains the Virgo Cluster and, near its edge, the Local Group. The latter is pulled toward constellation Virgo at a **Virgocentric peculiar velocity** of 168 km s⁻¹.

This is a minor perturbation of the much larger perturbation or inhomogeneity in the Hubble Flow from expansion. This perturbation (gravity anomaly) appears to arise from a “**Great Attractor**” toward constellation Centaurus (and located within the range of the Centaurus Supercluster⁵⁷⁷), much of which appears to be dark matter. However, the existence of this attractor is controversial, though it seems likely that there are perturbations in the Hubble Flow. (IMA2 p. 1070-1)

The **Shapley Supercluster** or **Shapley Concentration** of galaxies is another postulated contributor of some of the perturbation attributed to the Great Attractor. It has a mass of $\sim 10^{16} h^{-1} M_\odot$ and a distance of about 200 Mpc, so lies much further from the suspected location of the Great Attractor. (see image to right)

Other nearby superclusters include the Hydra-Centaurus Supercluster, the Perseus-Pisces Supercluster, and the Coma Supercluster.

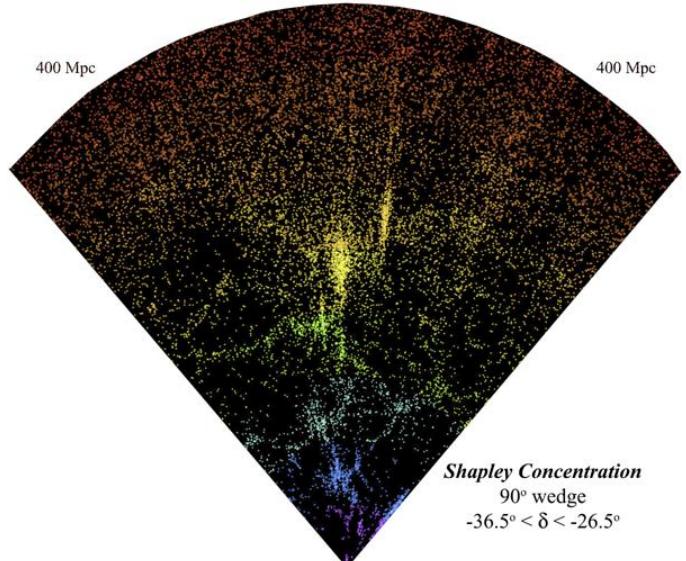
“The **CfA Homunculus** lies at the heart of the **Great Wall**, and the Coma Supercluster forms most of the homunculus structure”⁵⁷⁸ (see below).

Between the superclusters, which in large scale plots appear like clumpy **filaments**, are **voids** measuring up to 100 Mpc. Galaxies have not had time to form in these voids and then move out. The filaments appear connected, like the areas of intersection in a froth of soap bubbles.

The largest scale aggregations of clusters, seen in the 2dFGRS (2dF Galaxy Redshift Survey),⁵⁷⁹ Sloan Digital Sky Survey, and the Galaxy And Mass Assembly GAMA survey, are the **Great Wall** and the **Southern Wall**, the largest known discrete structures in the universe. (IMA2 p. 1074–1080, further details omitted.)

At the very largest scales, the cosmological principle (a sacred cow assumption) dictates that there should be no localized attractor that would impart a preferred direction in space.

The following near-Infrared panoramic views of galaxy clusters out to **z=0.1** (comoving radial distance of 400 Mpc) of the entire near-infrared sky “reveals the distribution of galaxies beyond the Milky Way. The image is derived from the 2MASS (Two Micron All Sky Survey) Extended Source Catalog (XSC)—more than 1.5 million galaxies, and the Point Source Catalog (PSC [also from 2MASS])—nearly 0.5 billion Milky Way stars. The galaxies are color coded by redshift obtained from the UGC, CfA, Tully NBGC, LCRS, 2dF, 6dFGS, and SDSS surveys (and from various observations compiled by the NASA Extragalactic Database), or photometrically deduced from the K band (2.2 um [near-infrared]). Blue are the nearest sources ($z < 0.01$); green are at moderate distances ($0.01 < z < 0.04$) and red are the most distant sources that 2MASS resolves ($0.04 < z < 0.1$). The map is projected with an equal area Aitoff [projection] in the Galactic system (Milky Way at center).”⁵⁸⁰ I have included versions without and with labels for the clusters. A more accurate description of the color coding is included just below, also from the Jarrett article. The article cited is also the source of the Shapley Concentration image, above.



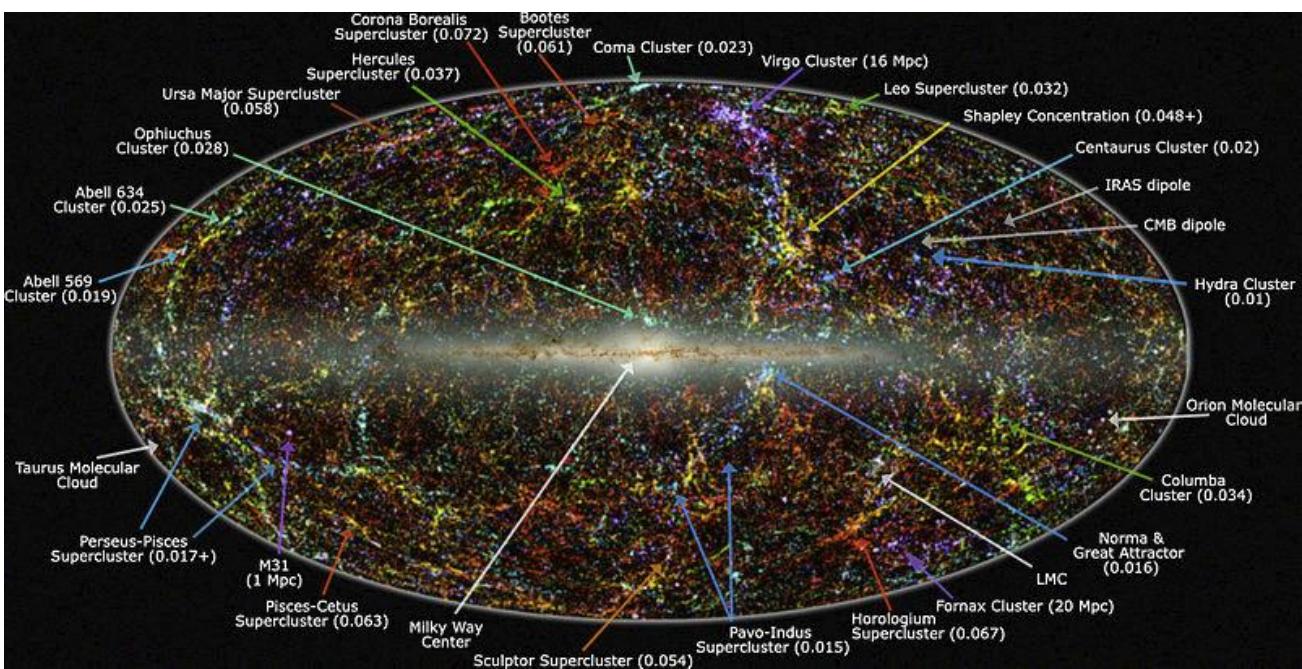
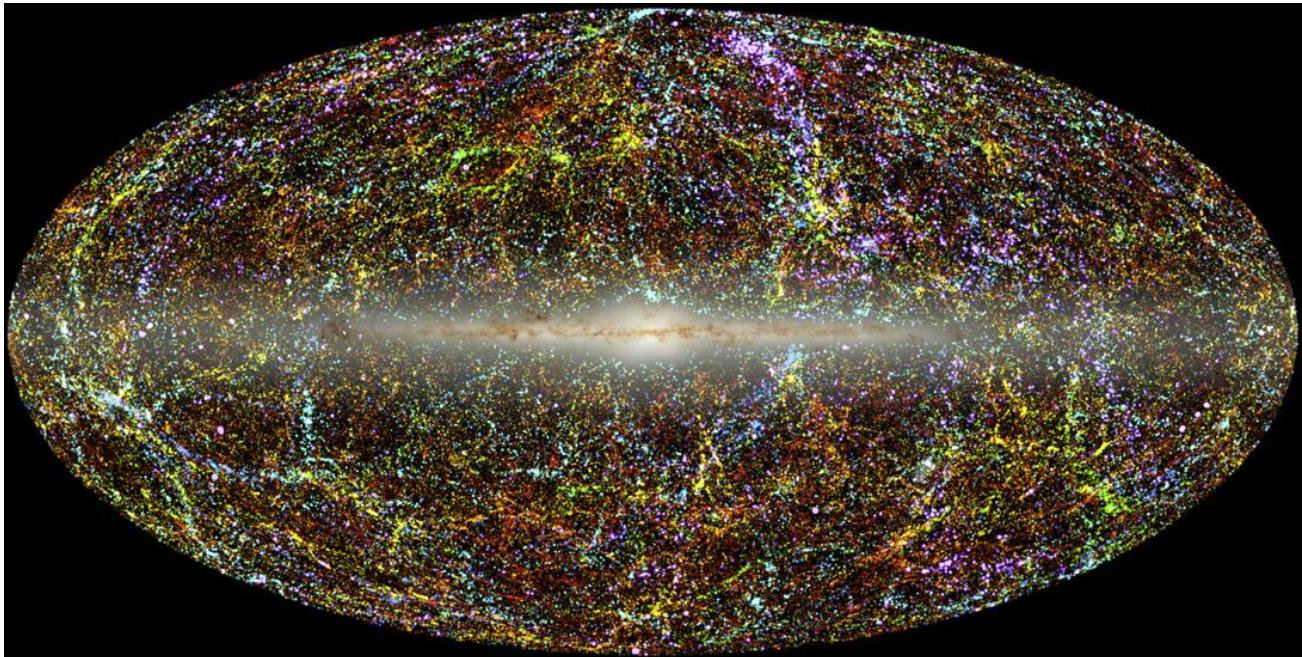
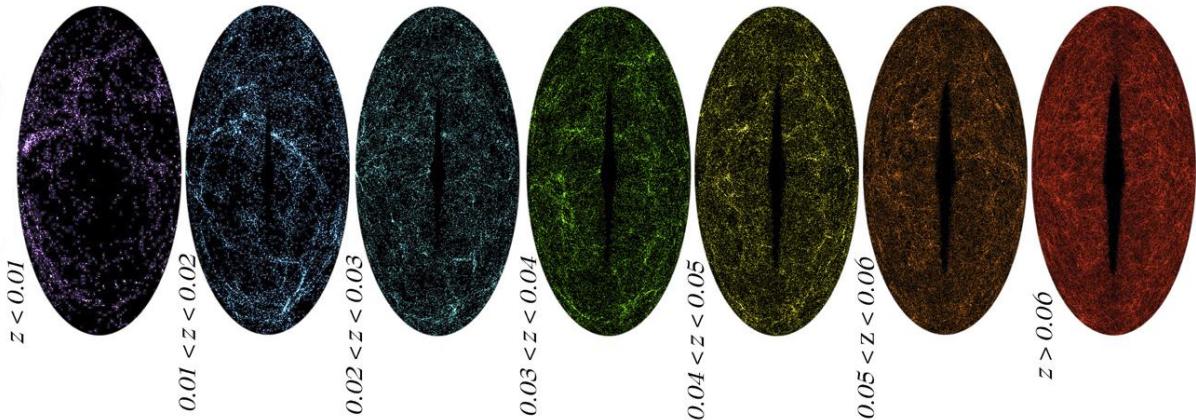
⁵⁷⁷ http://en.wikipedia.org/wiki/Great_Attractor

⁵⁷⁸ http://en.wikipedia.org/wiki/Galaxy_filament

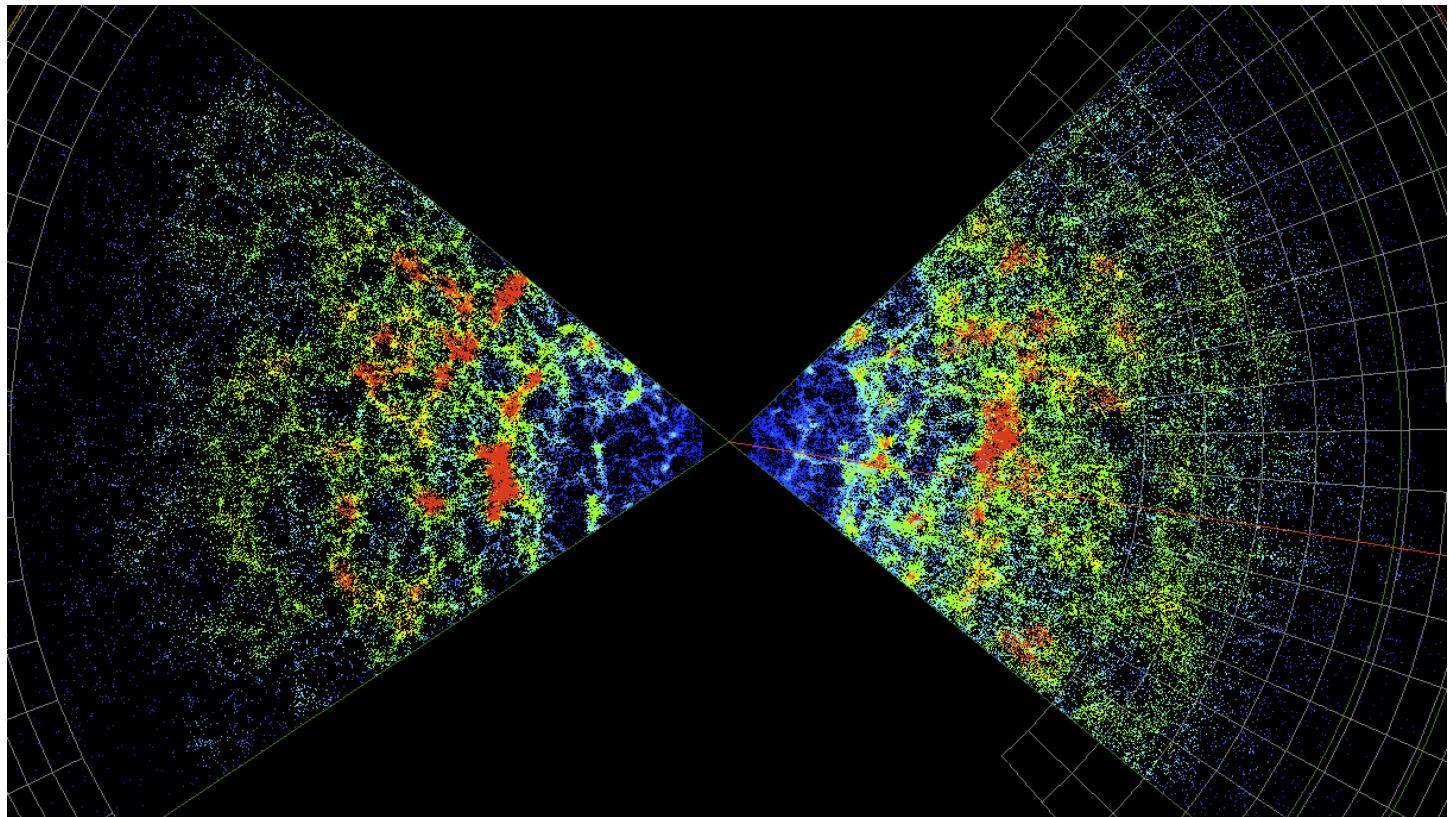
⁵⁷⁹ <http://www.mso.anu.edu.au/2dFGRS/>

⁵⁸⁰ <http://spider.ipac.caltech.edu/staff/jarrett/papers/LSS/> , published as

Jarrett TH, "Large Scale Structure in the Local Universe: The 2MASS Galaxy Catalog", 2004, PASA, 21, 396
Labeled image from http://en.wikipedia.org/wiki/Great_Attractor



Another galaxy survey, the **2dF Galaxy Redshift Survey (2dFGRS)**, provides us with the following distribution of clusters out to $z=0.25$:⁵⁸¹



The galaxies cover an area of approximately 1500 square degrees at $z<0.25$
selected from the extended APM Galaxy Survey in three regions:
an NGP [part-way toward the North Galactic Pole] strip,
an SGP [part way to the South Galactic Pole] strip,
and random fields scattered around the SGP strip.

⁵⁸¹ <http://magnum.anu.edu.au/~TDFgg/> (image cropped by MCM)
Page 225 of 249

Active Galaxies (Chapter 28)

This is a fascinating area of cutting-edge astronomy, and deserves far more extensive coverage than I have been able to give to it here.

Classification and Characteristics of Active Galaxies

An **active galaxy** is a galaxy hosting an unusually **active galactic nucleus (AGN)**. Active galaxies of a wide variety of types have been described and the terminology has been confusing. An AGN may be loosely defined as the central region of a galaxy which exhibits central luminosity (energy release) much higher than typical for stellar processes. Another definition, which ignores the degree of luminosity, might be, "A galaxy nucleus which shows evidence for accretion onto a supermassive black hole [SMBH]."⁵⁸² The presence of a SMBH is not alone sufficient—it must be actively accreting and emitting copious radiation of some form.

Active Galaxies are thought to have SMBHs in their nuclei. ZXI states (lecture 5) that most ellipticals and bulges contain SMBHs.

In addition, the following features may be found. (This body of knowledge has been confusing but is clarified by a unification theory, see later section):

Name or Class	Radio Q=Quiet L=Loud	Characteristics ⁵⁸³
Seyfert 1	RQ	Seyfert Galaxy 1: Moderate overall luminosity; strong/visible blue optical continuum; host galaxy clearly visible; broad + narrow lines; narrow lines have high ionization; weak radio emissions; moderately variable X-ray emissions; most are spiral galaxies; variable
Seyfert 2	RQ	Seyfert Galaxy 2: Similar to Seyfert 1; narrow emission lines but no broad lines; weaker or no blue optical continuum; weak radio; weak X-ray emissions (possibly due to absorption); most are spiral galaxies; not variable
QSR	RL	Radio-loud Quasar: Strong radio; strong optical continuum; broad + narrow emission lines; some polarization; FRII; variable
QSO	RQ	Radio-quiet Quasar (Quasi Stellar Object): Broad + narrow emission lines; host galaxy barely or not visible; strong optical continuum; weak polarization; variable
QSO-2	RQ	QSO-2: same as QSO but missing broad lines; not many currently known (some are IRAS-QSOs)
LINER	RQ/L	Low Ionization Nuclear Emission-line Region galaxy: Weak/no continuum; narrow lines of low ionization & moderate strength; can be either radio quiet or loud;
BLRG	RL	Broad Line Radio Galaxy: Similar to Seyfert 1 but strong radio; broad + narrow lines; strong radio emissions; FR II; weak polarization elliptical galaxies; variable

⁵⁸² http://www.astro.virginia.edu/class/whittle/astr553/Topic15/Lecture_15.html

⁵⁸³ Table adapted and revised by MCM from ibid. plus IMA2 p. 1107 etc.

NLRG	RL	Narrow Line Radio Galaxy: Similar to Seyfert 2 but strong radio; narrow emission lines only; strong radio emissions; FR1 and FRII, no polarization elliptical galaxies; not variable
PRG-II	RL	Powerful Radio Galaxy of Fanaroff-Riley class II (FRII): Edge brightened, powerful jet; unspecified optical spectrum (could be BLRG/NLRG/LINER)
RG-I	RL	Radio Galaxy of Fanaroff-Riley class I (FRI): Similar to PRG-II except lower radio luminosity & Fanaroff-Riley class I FRI (edge darkened, lower power jet)
LD-QSR	RL	Lobe Dominated Radio Loud Quasar: FR-II radio morphology; optically similar to QSO
CD-QSR	RL	Core Dominated Radio Loud Quasar: Flat spectrum; optically similar to QSO
BL-Lac	RL	BL Lacertae object: A type of Blazar . Strong featureless continuum, with no or weak emission lines, highly and rapidly variable; strong polarization; strong radio flat spectrum core; 90% are in ellipticals
OVV	RL	Optically Violently Variable Quasar: A type of Blazar . Similar to BL Lac but normal QSO spectrum and much more luminous than BL-Lacs Broad + narrow emission lines; strong radio; strong polarization; rapid variability.

Seyfert Galaxies

Seyfert galaxies (Types 1 and 2) are a class of galaxies with active galactic nuclei that produce spectral line emissions from highly ionized gas, first observed in 1908 by Edward Fath (1880-1959) and named after Carl Keenan Seyfert (1911-1960), who first described the class in 1943. The centers of Seyfert galaxies are thought to contain SMBH's with masses between 10^7 and $10^8 M_{\odot}$. "Seyfert galaxies are characterized by extremely bright nuclei, and spectra which have very bright emission lines of hydrogen, helium, nitrogen, and oxygen. These emission lines exhibit strong Doppler broadening, which implies velocities from 500 to 4000 km/s, and are believed to originate near an accretion disk surrounding the central black hole."⁵⁸⁴ The lines are from both allowed and forbidden transitions. Seyfert II have only narrow emission lines. "It is believed that Type 1 and Type 2 galaxies are in essence the same, and they only differ due to the angle at which they are observed. This is known as Seyfert Unification theory. In Type 2 Seyferts it is believed that the broad component is obscured by dust and/or by our viewing angle on the galaxy."⁵⁸⁵ Seyferts intermediate between 1 and 2 are classified as 1.5, etc. Seyferts show a featureless continuum devoid of lines [i.e., between what lines are present], contributing much of the luminosity and often overwhelming the light of the galaxy's stars (IMA2 p. 1086). See also table.

Examples include Mrk 1242 and Mrk 1157 (as listed in E. Benjamin Markarian's catalog) and NGC 4151 (the "Eye of Sauron", a nearby galaxy with readily discernible disk and bright nucleus).

⁵⁸⁴ http://en.wikipedia.org/wiki/Seyfert_galaxy

⁵⁸⁵ ibid.

Spectral Energy Distributions of Active Galactic Nuclei

The spectral energy distributions (SEDs, typically plotting flux density against frequency) of many types of AGNs including quasars exhibit varying degrees of the following (summary mostly from *IMA2* p. 1088).

- **Wide Range:** A very broad energy distribution spanning 8 to 10 or 15 orders of magnitude of photon frequency. (This is quite different from a pure Planck blackbody distribution of a star, which rises and falls over a much narrower range.) AGN spectra are thought to be a combination of thermal (blackbody, low polarization) and non-thermal emissions (synchrotron radiation, often polarized).
- **Radio:** Either relatively radio-loud or radio-quiet
- **IR:** A “**turnover**” (flattening off of the rise of the SED) at the “**infrared IR Bump**” The bump is a thermal feature, thought to be emission from warm dust grains with $T \leq 2000\text{K}$.
- **Light and UV:** A “**Big Blue Bump**” from high UV output (a thermal component contributing much of the bolometric luminosity of the source, perhaps from an optically thick accretion disk or free-free (Bremsstrahlung) emissions).
- **X-ray:** A wide ongoing continuum into X-ray and even Gamma-ray⁵⁸⁶ photon energies

The SEDs of AGNs therefore do not follow a simple power law, which would be expressed by

$$F_v \propto v^\alpha$$

where F is flux density for frequency v , and α = spectral index, but such a relationship may be useful to parameterize the continuum. Such power law behavior would be characteristic of synchrotron radiation generated by relativistic electrons circling magnetic field lines under the Lorentz force.

Here are representative SEDs for quasars of radio loud and radio quiet type (image annotated by MCM):⁵⁸⁷

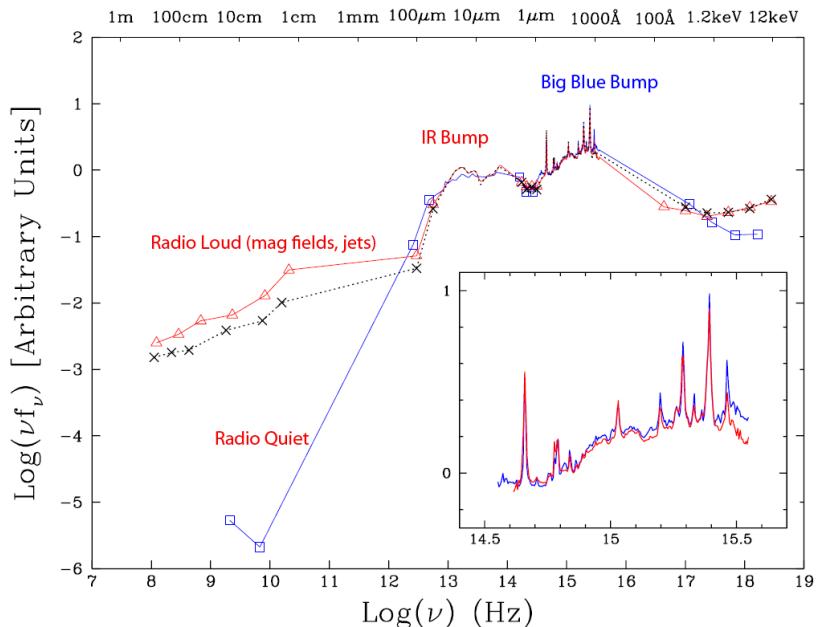


Fig. 7.— Median SEDs for the entire radio-loud sample (open triangle, red) and radio-quiet sample (open square, blue). Also shown is the median SED for only the low-redshift radio-loud objects (cross, dotted-line) for comparison (§5.3.1). The inset shows only the UV-optical region.

⁵⁸⁶ http://imagine.gsfc.nasa.gov/docs/science/know_11/active_galaxies.html

⁵⁸⁷ Zhaohui Shang, et al, “The New Generation Atlas of Quasar Spectral Energy Distributions from Radio to X-rays”, <http://arxiv.org/abs/1107.1855> (image annotated by MCM)
Page 228 of 249

Such SEDs over a large range of frequencies draw on data from many sources: VLA, Submillimeter array, FIRST, Spitzer, IRAS, ISO, 2MASS, SDSS, HST, FUSE, GALEX, ROSAT, Chandra, XMM, and Compton, etc. (ZXI lecture 7 and *here*⁵⁸⁸)

Radio Galaxies

The first found, **Cygnus A** (3C 405)⁵⁸⁹, was discovered by Grote Reber (1911-2002) in 1939. Karl Jansky (1905-1950) developed the early field of radio astronomy. Cyg A is the brightest extragalactic radio source and only exceeded in output by the Sun and Cassiopeia A in the MW (Sagittarius A* at the center of the MW is not strong). Baade and Minkowski (1895-1976) were able to locate an optical counterpart for Cyg A—it is a cD galaxy whose nucleus is surrounded by dust. Its distance is ~240 Mpc, so that the radio emission is extremely bright, the radio luminosity estimated by IMA2 p. 1092 at $L_{\text{radio}} = 4.8 \times 10^{37}$ W. (See image below.)

Radio Galaxies have also been divided into **Broad Line Radio Galaxy** (BLRGs, similar to Seyfert 1s but radio loud) and **Narrow Line Radio Galaxy** (NLRG, similar to Seyfert 2 but radio-loud, giant or supergiant ellipticals, include Cyg A).

One of the largest radio galaxies is 3C 236, spanning about 4.5 Mpc—it is a FR II radio galaxy.⁵⁹⁰

A nearby example is Centaurus A (NGC 5128), which produces a prominent radio- and X-ray-emitting jet. The galaxy is hard to categorize—it is listed by NED as a “peculiar” S0 lenticular galaxy and by Simbad as a Seyfert 2 galaxy. (See image below.)

Radio Lobes and Jets

Radio galaxies may emit radio from extended lobes fed by **jets** or from a compact core or a halo. The jets tend to be strongly collimated (see model discussion below). At least half of the stronger radio sources have observable jets.

The most powerful radio galaxies often appear to have one-sided jets, but this is often an effect of relativistic beaming (“headlight” phenomenon) or extinction of the counterjet.

High resolution imaging of the jet regions, e.g., of NGC 6251, shows that the jet can be traced back to the core of the galaxy. Jets often display **knots** or **hot spots** of enhanced luminosity (in radio, visible, or X-ray) suggesting variable outputs of material. M87 (Messier 87, aka Virgo A) is one of the brightest radio objects in the sky, and exhibits a highly collimated jet with knots, extending out 1.5 kpc from the core of its supergiant elliptical galaxy.⁵⁹¹ The X-ray luminosity of M87 overall is 50 times the radio luminosity.

Jets do not always remain straight, and may appear curved or “windblown”, apparently distorted by the galaxy’s movement causing friction with its intracluster gas. See for example, NGC 1265 which is moving through the Perseus cluster (Abell 426).⁵⁹²

⁵⁸⁸ Zhaohui Shang, 2009, <http://202.127.29.4/cosmology09/talk/ShangZhaohui.ppt>

⁵⁸⁹ The 3C catalog is *Third Cambridge Catalogue of Radio Sources*, http://en.wikipedia.org/wiki/3C_Catalog

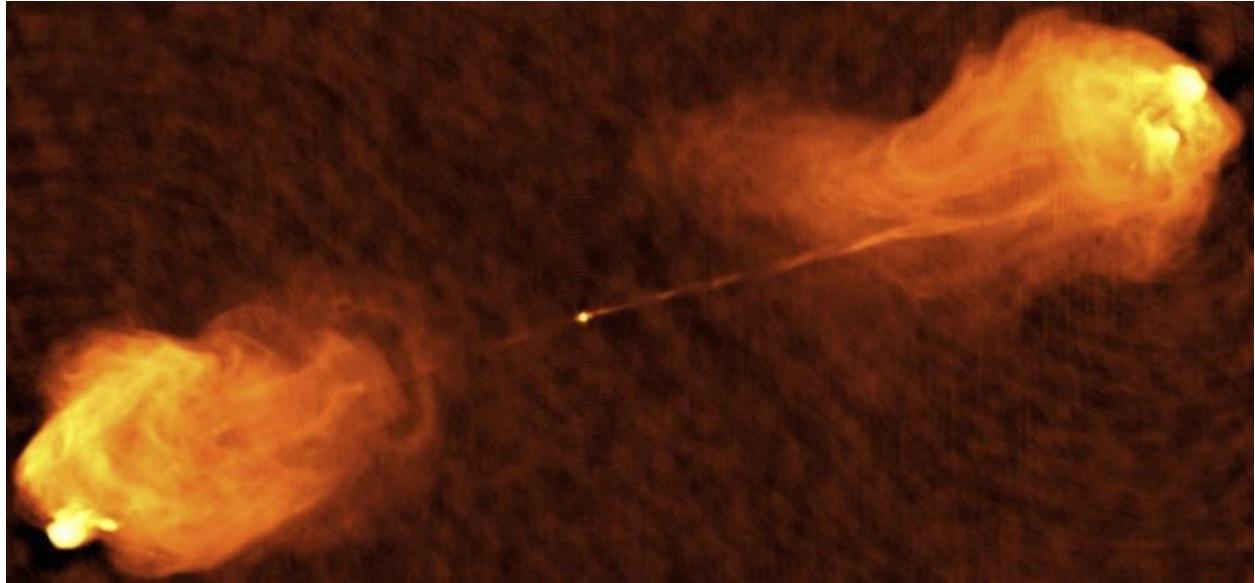
⁵⁹⁰ Tremblay GR; et al; “Episodic Star Formation Coupled to Reignition of Radio Activity in 3C 236”, *ApJ* Volume 715, Issue 1, pp. 172-185 (2010)

⁵⁹¹ http://en.wikipedia.org/wiki/Messier_87

⁵⁹² Abell = *Abell catalog of rich clusters of galaxies*, http://en.wikipedia.org/wiki/Abell_catalogue



Centaurus A (NGC 5128), superimposed dual radio lobes (VLA) and optical (probably HST),
Dark dust lanes partially obscure the optical elliptical galaxy.
Distance is 10 Mly, image spans 18.2 arcmin, 1998⁵⁹³



Cygnus A, showing central active galaxy (bright dot), collimated highly asymmetrical jets, and radio lobes.
5 GHz, 0.5" resolution, 0.03° FOV, redshift $z = 0.057$, $d = 230$ Mpc. 1983 (image: NRAO/AUI)⁵⁹⁴

⁵⁹³ <http://hubblesite.org/newscenter/archive/releases/1998/14/image/1/>

⁵⁹⁴ <http://images.nrao.edu/110>

Quasars

Much of what follows under this section applies to galaxies with AGNs, whether or not they are classical radio-loud quasars.

A **quasi-stellar radio source (quasar)** is a very energetic and distant active galactic nucleus. Quasars are extremely luminous and often highly redshifted sources of EM energy, including radio waves and visible light. They appear point-like, similar to stars, rather than as more extended galaxy-like sources or radio lobes. Quasars are all believed to have a SMBH at the center. The size of a quasar is 10–10,000 times the Schwarzschild radius of the black hole. The quasar's emissions arise from the accretion disc.

The term *quasar* has undergone evolution and can be ambiguous. It originally described radio-loud sources (now usually abbreviated **QSRs**), but *quasar* has been extended to include **radio-quiet quasars (quasi-stellar objects**, now abbreviated **QSOs**). The name and abbreviation *quasi-stellar objects* or *QSO* are also used to indicate quasars of either type.

The first observed in optical was by Allan Sandage and Thomas Matthews in 1960. They were searching for a counterpart to the radio source **3C 48**. This faint quasi-stellar object was emitting more intense radio waves and ultraviolet radiation than a typical star. Its SED revealed a unique spectrum displaying broad emission lines that could not be identified with any known element. A similar radio-loud quasar was found for **3C 273**, which happened to have a prominent jet and a redshift $z = 0.16$, making spectroscopic identification of lines difficult. The absolute visual magnitude M_V was calculated as -26.2 , yielding a visual luminosity of 10^{39} W. The radio luminosity was about 7×10^{36} W and the bolometric luminosity is estimated at 2.2×10^{40} W.⁵⁹⁵ The observed range of bolometric luminosities for quasars is 10^{38} to 10^{41} W (*IMA2* p. 1097). The most energetic quasars have 10^5 times the bolometric luminosity of normal galaxies like the Milky Way.

The redshift of 3C 48 was subsequently found to be $z = 0.367$, corresponding to a radial speed of $0.303c$ and a Hubble distance of $900h^{-1}$ Mpc, making it one of the most distant of known objects.

The most distant quasar yet discovered by the UKIDSS Large Area Survey ULAS⁵⁹⁶ is **ULAS J1120+0641** with a redshift $z = 7.085$, thus seen only some 770 million years after the Big Bang. “It took 12.9 billion years for its light to reach us.”⁵⁹⁷ This redshift corresponds to a comoving or proper distance of 29 billion lightyears.⁵⁹⁸ The size of the black hole is estimated at about $2 \times 10^9 M_\odot$. (An even further object is UDFj-39546284, a compact galaxy of blue stars with unconfirmed $z = \sim 10$, corresponding to a comoving distance ~ 32 billion lightyears.⁵⁹⁹

The high luminosity of quasars we observe may be in part due to the directionality of the jets. The emission may not be uniform for all angles, and the quasars that we observe may be those with jets pointing toward us. Highly luminous quasars and AGN's overwhelm observation of the surrounding galaxy, but lower-luminosity AGNs appear like galaxies (ZXI lecture 7).

Quasar Spectra

These exhibit a very broad energy distribution spanning 8 to 10 or 15 orders of magnitude, dominated at low frequencies by the jet radiation. A turnover (flattening off of the rise) and infrared bump occurs in far infrared. Peak power is radiated in the low-energy gamma rays. A prominent blue bump is often seen due to excess of UV output (at $10^{14} - 10^{16}$ MHz), and quasars visually appear quite blue. There are often absorption lines, often Doppler broadened and narrow lines... (see *IMA2* p. 1098 for further details).

The following graph (in a 2001 paper for which ZXI is a co-author) is a composite of over 2200 SDSS spectra of quasars with at least one broad emission line, adjusted to a rest frame. Broad absorption line quasars (BAL quasars) are not included. The spectrum shows broad and narrow lines (the former are not seen in Type

⁵⁹⁵ http://ned.ipac.caltech.edu/level5/March02/Courvoisier/Cour3_5.html

⁵⁹⁶ UKIRT=United Kingdom Infrared Telescope (on Mauna Kea); UKIDSS = UKIRT Infrared Deep Sky Survey; ULAS = UKIDSS Large Area Survey ULAS

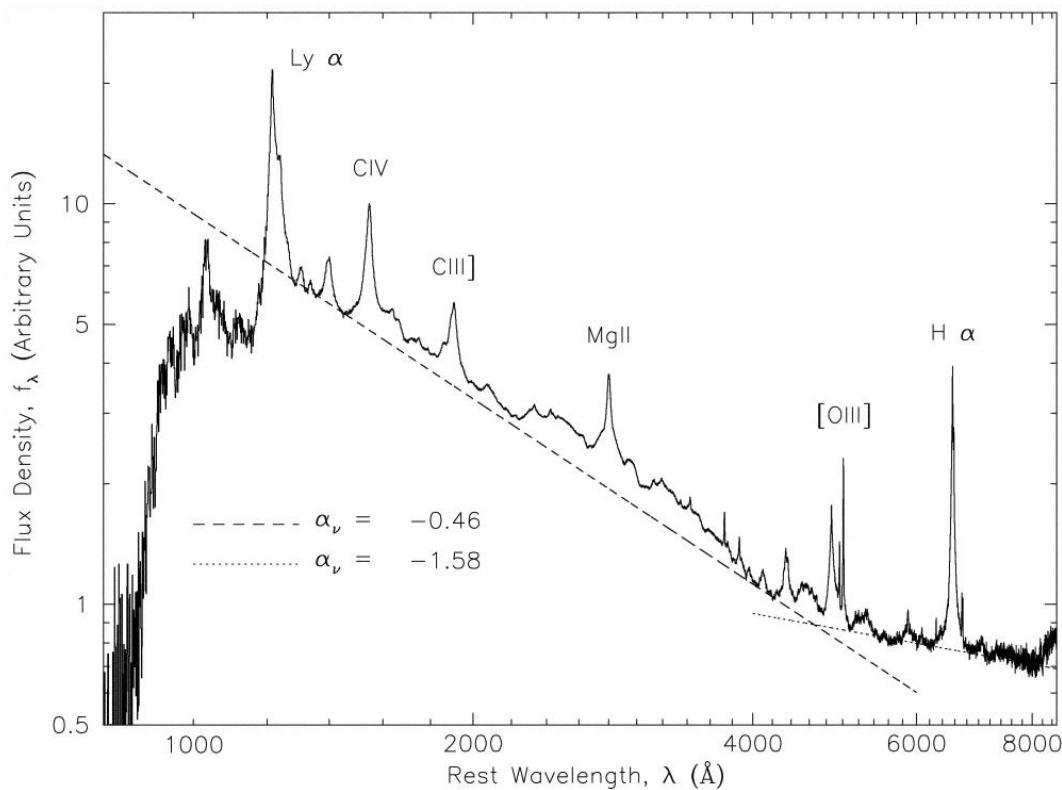
⁵⁹⁷ <http://www.eso.org/public/news/eso1122/>

⁵⁹⁸ http://en.wikipedia.org/wiki/ULAS_J1120%2B0641

⁵⁹⁹ UDFj-39546284 distant compact galaxy:

- <http://www.astro.ucla.edu/~wright/CosmoCalc.html> (Cosmology Calculator)
- <http://en.wikipedia.org/wiki/UDFj-39546284>

2 AGNs, but there is some overlap). “At wavelengths longer than the Ly α emission, the continuum of the geometric mean composite is well fitted by two power laws, with a break at $\approx 5000 \text{ \AA}$. The frequency power-law index, α_ν , is -0.44 from ≈ 1300 to 5000 \AA and -2.45 redward of $\approx 5000 \text{ \AA}$. The abrupt change in slope can be accounted for partly by host-galaxy contamination at low redshift.”⁶⁰⁰



As with Seyferts and AGNs in general, it is believed that for Quasars of Type 1 that have broad emission lines, we can see their high velocity gas (presumably in the accretion disk) whereas Type 2 quasars (which lack broad emission lines) are tilted so that we do not see the high-velocity gas (or simply lack it (ZXI lecture 7).

Ultraluminous Infrared Galaxies

Many of these were found by the 1983 Infrared Astronomical Satellite (IRAS) and by later infrared-sensing satellites. A **Luminous Infrared Galaxy (LIRG)**, is a galactic body whose defining characteristic is that it emits more than $10^{11} L_\odot$ in the far-infrared. An even more luminous system, emitting more than $10^{12} L_\odot$ in the far-infrared, is called an **Ultraluminous Infrared Galaxy (ULIRG)**. Most LIRGs and ULIRGs emit at least 90% of their light in the infrared. The infrared may arise from absorption and re-radiation by dust enshrouding the radiation from an active galactic nucleus or quasar. These galaxies are cataloged in the IRAS Revised Bright Galaxy Sample (RBGS).⁶⁰¹

High Cosmological Redshift of Quasars

As previously noted, high redshifts are seen with quasars, as high as 7.1. In the SDSS survey, currently identifying more than 60,000 quasars, those in the top 1000 in redshift have z values ranging from 4.1 to 6.0.⁶⁰² There are more than 1000 having $z > 4$. A redshift of $z=6.0$ corresponds to an age at emission of 2.2

⁶⁰⁰ Vanden Berk DE, et al, “Composite Quasar Spectra From The Sloan Digital Sky Survey”, *The Astronomical Journal*, 122:549-564, 2001 August.

⁶⁰¹ <http://arxiv.org/abs/astro-ph/0306263>

⁶⁰² <http://cas.sdss.org/dr7/en/proj/advanced/quasars/query.asp>

Note: for sorted output, add at the bottom of the suggested query order by z desc

Gigayear (i.e., the age relative to the Big Bang, which occurred 13.7 Gyr ago), a light travel time of 11.48 Gyr,⁶⁰³ and a comoving radial distance of 21.07 Gly.⁶⁰⁴

Redshifts can be caused by ordinary Doppler effects and by gravitational effects on light moving out of a strong gravity field.⁶⁰⁴ In addition, **Cosmological Redshifts** are caused by the expansion of space through which the light travels. The total elongation of wavelength depends on how much expansion has occurred. For such great distance, it is customary to quote z rather than an inferred distance, computed by wavelength in the emitter's rest frame compared to wavelength in the observer's rest frame.

$$z = \frac{\lambda_{\text{obsv}} - \lambda_{\text{emit}}}{\lambda_{\text{emit}}} = \frac{R_{\text{obsv}} - R_{\text{emit}}}{R_{\text{emit}}}$$

or

$$\frac{R_{\text{obsv}}}{R_{\text{emit}}} = 1 + z$$

where R is the size of the universe (to be later defined). Thus a redshift of 7 means that the universe is 8 times larger than when the light was emitted (*IMA2* p. 1099–1100).

Quasar Evolution

Quasars at large redshifts appear to us as they did many billions of years earlier. Equivalently, we see further into the past when viewing quasars at increasingly large redshifts. Bright quasars were more common (had greater space number density) in the distant past (large redshift). To remove the confounding effect of the expansion of space in comparing density assessments, we define the **comoving space density** as:

$$\text{comoving space density} = \frac{\text{Number of objects at redshift } z \text{ per Mpc}^3}{(1+z)^3}$$

This is a way of saying the density the object will have at the current time, for which $z = 0$. As long as objects are not being created or destroyed or evolving to other categories, etc., their comoving space density is constant despite the expansion of space.

The comoving space density of quasars brighter than $M_B = -25.9$ is 1000 times higher at $z=2$ than at $z=0$. Quasar number density peaks at about $z = 2.2$, when the universe was about 1/4 of its present age, and that 95% of quasars are found between 5% and 50% of the current universe's age (*ZXI lecture 7*).⁶⁰⁵ However, this appears to be due to a shift with time of the luminosity curves $\Phi(M_g)$ rather than of the number of quasars⁶⁰⁶ (see also *IMA2* p. 1101 for diagram). Assessment of quasar populations also suggests that quasars on the average tend to be dimmer for lower redshifts near 0 and thus greater proximity compared to quasars with $z=2$.

But for $2 < z < 3$, the comoving space densities for AGNs in X-ray wavelengths are seen (Silverman 2005) to peak at about $z = 2.5$ and to falloff for higher z values:⁶⁰⁷

⁶⁰³ <http://www.astro.ucla.edu/~wright/CosmoCalc.html>

⁶⁰⁴ http://en.wikipedia.org/wiki/Gravitational_redshift

⁶⁰⁵ <http://www.astro.ucla.edu/~wright/CosmoCalc.html> (to calculate age by given redshift)

⁶⁰⁶ Richards GT et al, “The 2dF-SDSS LRG and QSO (2SLAQ) Survey: the $z < 2.1$ quasar luminosity function from 5645 quasars to $g = 21.85$ ”, *Monthly Notices of the Royal Astronomical Society (MNRAS)*, Volume 360, Issue 3, pp. 839–852 07/2005

⁶⁰⁷ Silverman JD, et al, “Comoving Space Density of X-ray-selected Active Galactic Nuclei”, *ApJ* 624 630, 2005
Page 233 of 249

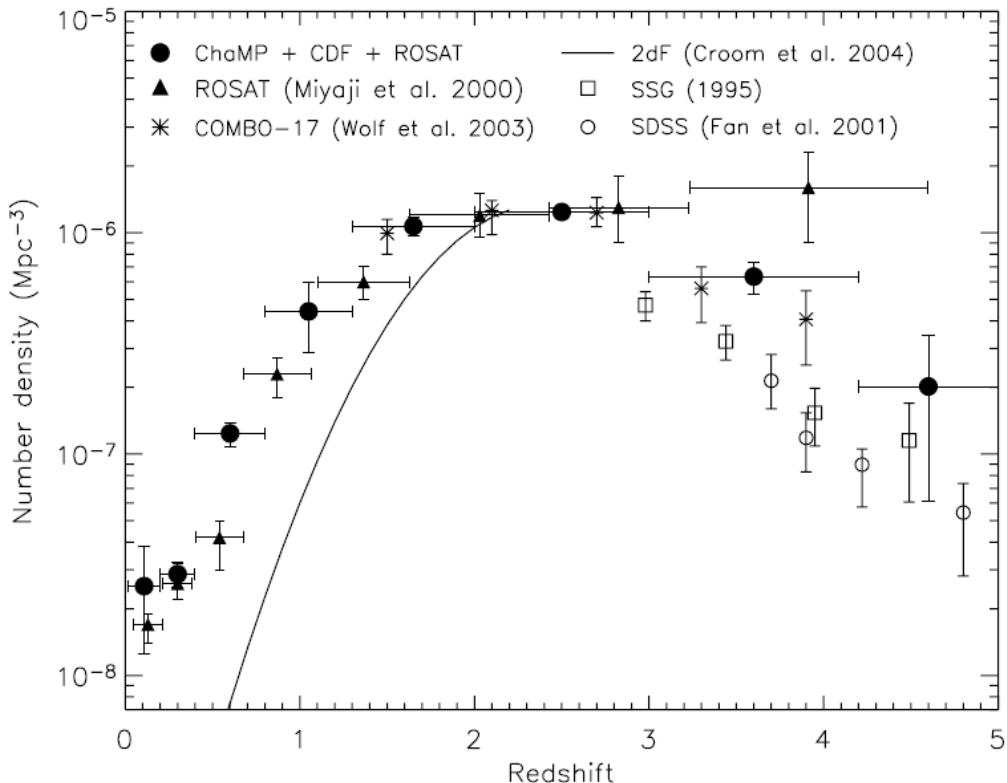


FIG. 9.—Comoving space density of 217 *Chandra + ROSAT* AGNs selected in the soft (0.5–2.0 keV) band with $\log L_X > 44.5$ compared to the optical surveys. The optical space densities have been scaled to match the X-ray points at $z = 2.5$ for ease of comparison.

Perhaps the decreasing numbers of quasars at even higher z ($z > 3$) indicates that time for a growth phase of SMBHs must elapse before quasars can develop.

The existence of a quasar may be limited to a galactic dynamical timescale. For all radio objects studied in a HST study⁶⁰⁸ (10 RLQs=radio-loud quasars, 13 RQQs=radio-quiet quasars, and 10 RGs=radio galaxies) for which nuclear luminosities were brighter than absolute visual magnitude $M_V = -23.5$, the host galaxies were found to be massive ellipticals. (A few of the lower luminosity quasars were found to be in disk/bulge systems.) Other than having an AGN, the ellipticals appeared indistinguishable from lower- z quiescent galaxies found in cluster centers. All of the RLQs had SMBHs of at least $10^9 M_\odot$, and the RQQs had $M > 5 \times 10^8 M_\odot$. The greater abundance of RQQs compared to RLQs probably reflects the large size of the SMBH required for a RLQ.

It is possible that the strongest quasar radio sources require not only large SMBHs but also rotation of the SMBH (IMA2 p. 1104).

A 2004 SDSS-based study⁶⁰⁹ found that for 12,698 quasars with $0.01 \leq z \leq 2.1$, the majority of black holes with mass $> 10^{8.5} M_\odot$ (and therefore capable of powering quasars) were in place by $z \approx 2$. The quasar bolometric luminosities steadily increase with z from $z \sim 0.2$ to $z \sim 2.0$. The ratios of L_{bol} / L_{Edd} were no greater than 1, suggesting that Eddington luminosity “is still a relevant physical limit to the accretion rate of luminous broad-line quasars at $z \leq 2$.” Regarding quasar lifetimes, McLure and Dunlop state, “A fairly robust limit on the lifetime of quasars with black hole masses $\geq 10^{9.5} M_\odot$ is found to be $t_Q > 2 \times 10^8$ years.”

⁶⁰⁸ Dunlop JS, McLure RJ, et al, “Quasars, their host galaxies and their central black holes”, *Monthly Notices of the Royal Astronomical Society*, Volume 340, Issue 4, 9 APR 2003

⁶⁰⁹ McLure RJ, Dunlop JS, “Quasars, their host galaxies and their central black holes”, *Monthly Notices of the Royal Astronomical Society*, Volume 352 p. 1890, 2004

Timescale of AGN Variability

Many AGNs have luminosities of broad emission lines that can change by a factor of 2 in time intervals as short as a few days—this variability lags somewhat behind the variation in the continuum spectrum. (Narrow lines do not exhibit nearly this degree of variability.) X-ray and visible output can fluctuate a few percent over intervals as short as a few minutes. Some variations, such as the apparent magnitude of 3C 279, vary over much longer intervals up to several years. (*IMA2* p. 1104). See below regarding Blazars, BL Lacs and OVV Quasars.

“It is generally assumed that most, if not all, AGN are variable at some level. For example, virtually all of the AGN observed in the Hubble Deep Field were determined to exhibit detectable variability over a period of two years....

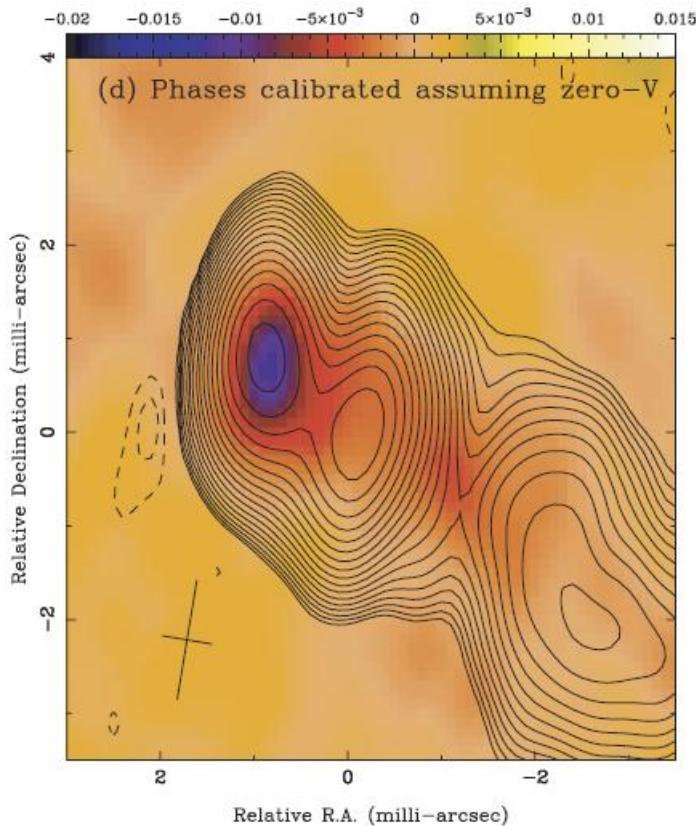
AGN Variability is an active area of research.

Polarization of Emissions

Quasar emission usually have low polarization of < 3% linear polarization. The **core-dominant radio sources** (aka **compact sources**) are the exceptions, having up to 35% linear polarization. The **lobe-dominant radio sources** have even higher polarization of up to 60% linear polarization (see for example 3C 47, shown on *IMA2* p. 1106). AGN **jets** have high polarization of up to 40 to 50%.

Unlike thermal radiation, which is unpolarized, **synchrotron radiation** from relativistic electrons circling in a magnetic field is polarized. This mechanism may account for some of the polarization seen with quasar emissions.

In addition to linear polarization, circular polarization may be observed in jets. It can arise intrinsically in the synchrotron process, or be produced secondarily by Faraday rotation:



Circular Polarization Distribution⁶¹⁰ (maximal in core) of Part of Jet of 3C 273
Circular polarization (color-scale given in mJy beam⁻¹) superimposed on Stokes I contours.
The I contours for both sources start at 10 mJy beam⁻¹. Image date August 28, 2003.
“The additional phase calibration [for this image] has corrected the phase error and reveals
that the real circular polarization in the source is on the jet core and not on the jet
component...”

Fanaroff-Riley Luminosity Classes

“It was first noticed by B.L. Fanaroff and J.M. Riley (1974) that the relative positions of regions of high and low surface brightness in the lobes of extragalactic radio sources are correlated with their radio luminosity. This conclusion was based on a set of 57 radio galaxies and quasars... Fanaroff and Riley divided this sample into two classes using the ratio R_{FR} of the distance between the regions of highest surface brightness on opposite sides of the central galaxy or quasar, to the total extent of the source up to the lowest brightness contour in the map. Sources with $R_{FR} < 0.5$ were placed in Class I [like NGC 1265 and 3C 449], and sources with $R_{FR} > 0.5$ in Class II [like Cyg A, 3C 47, and most quasars]. It was found that nearly all sources with luminosity $L_{(178 \text{ MHz})} < \sim 2 \times 10^{25} \text{ h}_{100}^{-2} \text{ W Hz}^{-1} \text{ str}^{-1}$ were of Class I, while the brighter sources were nearly all of Class II... [IMA2 specifies this distinction for 1.4 GHz]

Fanaroff-Riley Class I (FR-I): Sources in this class have their low brightness regions further from the central galaxy or quasar than their high brightness regions. The sources become fainter as one approaches the outer extremities of the lobes and the spectra here are the steepest, indicating that the radiating particles have aged the most. Jets are detected in 80 per cent of FR-I galaxies.... FR-I sources are associated with bright, large galaxies (D or cD) that have a flatter light distribution than an average elliptical galaxy and are often located in rich clusters with extreme X-ray emitting gas...

Fanaroff-Riley Class II (FR-II): This class comprises luminous radio sources with hotspots in their lobes at distances from the center which are such that $R_{FR} > 0.5$. These sources are called edge-darkened, which was particularly apt terminology when the angular resolution and dynamic range used in observing the classical sources was not always good enough to reveal the hotspots as distinct structures. In keeping with the overall high luminosity of this type of source, the cores and jets in them are also brighter than those in FR-I galaxies in absolute terms; but relative to the lobes these features are much fainter in FR-II galaxies. Jets are detected in < 10 per cent of luminous radio galaxies, but in nearly all quasars. The jets have small opening angles (< 4 deg) and are knotty; the jet magnetic field is predominantly parallel to the jet axis except in the knots, where the perpendicular component is dominant... The most striking feature of the jets in the FR-II class is that they are often one-sided... FR-II sources are generally associated with galaxies that appear normal, except that they have nuclear and extended emission line regions. The galaxies are giant ellipticals, but not first-ranked cluster galaxies...”⁶¹¹

Blazars: BL Lacs and OVV Quasars

“Some AGNs known as **BL Lacs** (also referred to as **blazars**) are among the most active variable sources known in the universe.”⁶¹²

BL Lacs are named after the most well-known object, BL Lacertae or BL Lac.

“Blazars are thought to be active galactic nuclei, with relativistic jets oriented close to the line of sight with the observer.”⁶¹³ Collimated beams exhibiting slight fluctuation in the direction of the beam could cause marked variability in the radiation reaching the observer. BL Lacs have a featureless continuum, and rapidly vary over hours to days or months in brightness. They are strongly polarized and have few emission lines.

⁶¹⁰ Homan DC et al, “MOJAVE: Monitoring of jets in active galactic nuclei with VLBA experiments. II. First-epoch 15 GHz circular polarization results”, *The Astronomical Journal*, 2006, available at http://docs.lib.psu.edu/cgi/viewcontent.cgi?article=1681&context=physics_articles

⁶¹¹ http://ned.ipac.caltech.edu/level5/Glossary/Essay_fanaroff.html

⁶¹² http://gtn.sonomal.edu/resources/agn_variability/agn_variability.php

⁶¹³ <http://en.wikipedia.org/wiki/Blazar>

Redshifts where observed are high, indicating that they are at cosmological distances. Most are in elliptical galaxies.

Another subdivision of blazars are the **Optically Violent Variable quasars (OVV)**. These are like BL Lacs except that they are much more luminous and their spectra may display broad emission lines.

LINERS

These are **Low Ionization Nuclear Emission-line Region galaxies**: They have very weak or no continuum (low luminosities); but fairly strong emission lines of low-ionization species. They are similar to the low limit of the range of luminosity for Seyfert 2 class galaxies, and are detectable in starburst galaxies. It is unclear if they represent the low-luminosity limit of AGNs. (IMA2 p. 1107)

A Unified Model of Active Galactic Nuclei (including Quasars)

It is reasonable to pursue an attempt to simplify and unify the seemingly disparate types of entities with AGNs. For one, the Luminosity $L_{\text{H}\alpha}$ of AGNs in the H α line has been shown to be proportional to L_{FC} , the luminosity of the featureless continuum at 480 nm, for quasars, Seyferts, and NLRGs (IMA2 p. 1109). This suggests that the ionized H atoms are generated by photoionization by the continuum radiation. There is also a line of evidence that suggests that within Seyfert 2's are Seyfert 1 nuclei that are hidden by optically thick material, and that the light we see arrives indirectly by reflection.

Size of the AGN

The rapid time variability is the strongest clue as to the scale of the AGN. The maximum radius R of an object that can brighten throughout over a period of Δt is given by $R = c\Delta t/\gamma$, where γ is the Lorentz factor

$$\gamma = \sqrt{1 - v^2/c^2}$$

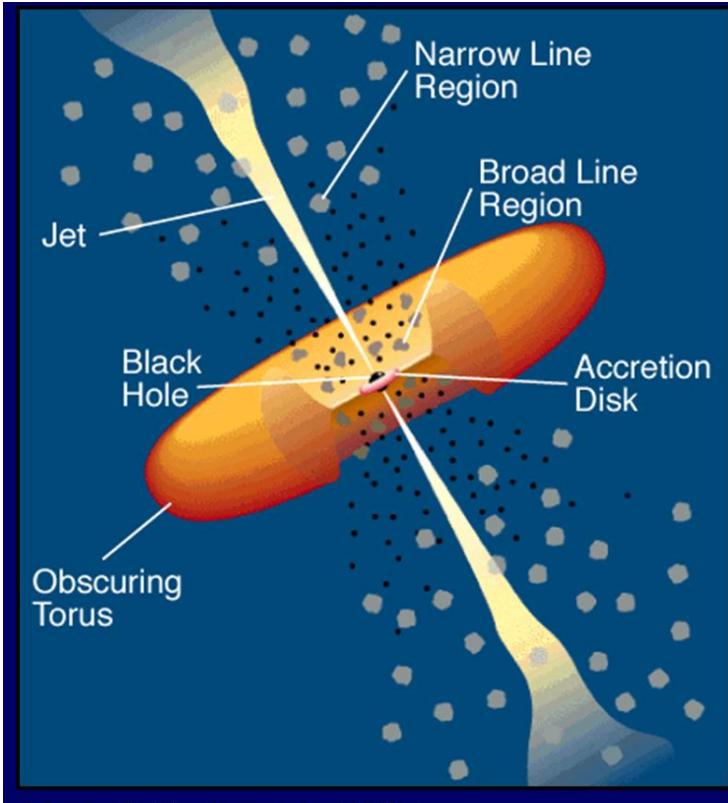
This is of course based on speed of light and special relativity considerations. For $\Delta t=1$ hour, $R = 7.2$ AU, a comparatively very small size for an object of such intense luminosity as an AGN. The Eddington limit sets an upper limit on luminosity, L_{Edd} , such that $L_{\text{AGN}} < L_{\text{Edd}}$. The lower limit for the AGN mass becomes $M_{\text{AGN}} > 3.3 \times 10^8 M_{\odot}$. The upper limit set by the Schwarzschild radius is $M = 3.7 \times 10^8 M_{\odot}$. The close convergence of these values supports that idea that a SMBH is present in the center.

Luminosity Through Accretion

The most efficient way to convert gravitational potential energy to other forms (heat, light) is by mass accretion. Accretion of matter onto a neutron star of $1.4 M_{\odot}$ releases 21% of the rest energy of the matter. With a BH, there is no surface to strike, the matter will just disappear through the event horizon. However, matter spiraling in decreasing orbit toward a black hole will pile up on an accretion disk, allowing viscous dissipation of PE and KE as heat and radiation. The smallest stable orbit is at $3R_s$, where R_s is the Schwarzschild radius. To reach this point, the matter must give up 5.2% of its rest mass in binding energy (the energy needed to remove it to infinity). For rapidly rotating black holes, the smallest stable orbit is at prograde 0.5 R_s (or retrograde 4.5 R_s), yielding up 42.3% of rest mass energy! "The accretion of matter through a disk around a rapidly rotating black hole is an extremely efficient way of producing large amounts of energy" (IMA2 p. 1111). The matter at the inner edge of the disk rotates at the same speed as the BH. The R_s increases for rising M_{BH} , so that characteristic disk temperature decreases for increasing M_{BH} , given by $T_{\text{disk}} \propto M^{-1/4}$. According to an example shown, a typical disk temperature could be 73000 K, which by Wien Displacement places the peak emission in the extreme UV, 39.7 nm (IMA2 p. 1112). More sophisticated calculations yield a disk temp of several hundred thousand K. The big blue bump would be a thermal signature from the UV excess arising at the disk.

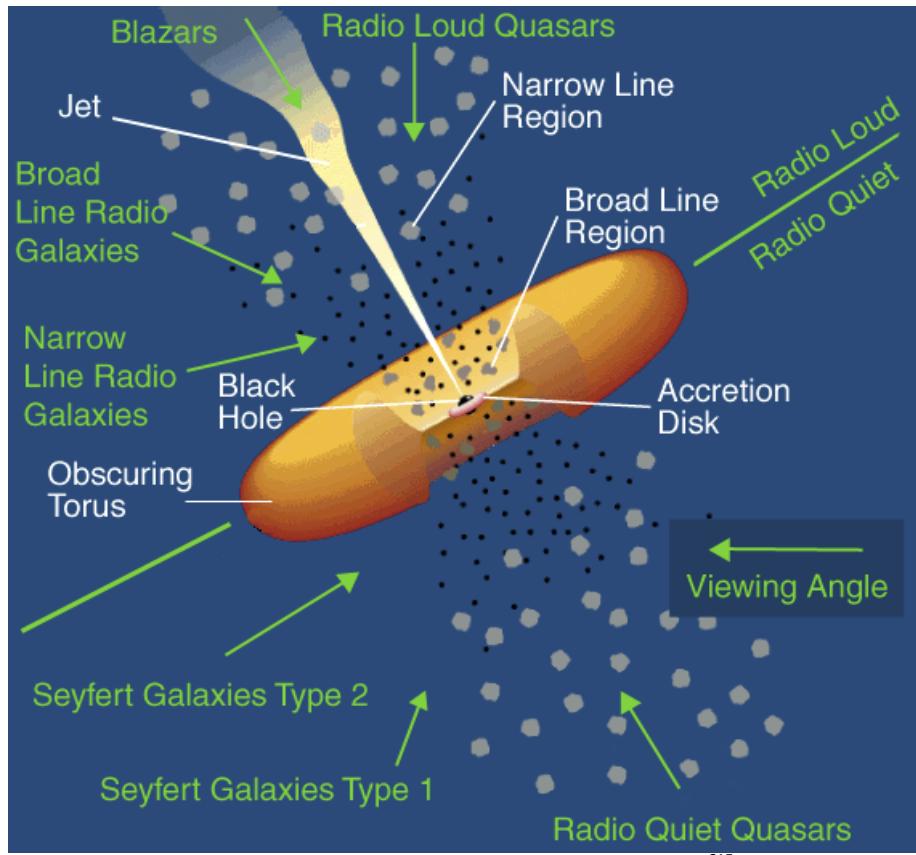
The Accretion Disk, Torus, and Size of the SMBH

These have been extensively modeled, and this can only be a limited summary. According to the unified model, the different appearances are due to differing viewing angles (in addition to differences in size and age, etc.) Here is one sample representation, though simplified compared to the textbook description and drawing:



Urry & Padovani, 1995
Unified Model of AGN with Black Hole and Accretion Disk.⁶¹⁴

⁶¹⁴ modified from www.ifa.hawaii.edu/~treister/talks/ifa08.pptx



Unified Model of AGNs (Annotated, from NASA)⁶¹⁵

If the disk overheats, it puffs up into an *ion torus*, supported by pressure of the hot ions. If the luminosity becomes super-Eddington, the radiation pressure lifts the matter into a more inflated *radiation torus*.

The disk may have several sections. The BH is in the center, unseen and tiny compared to the size of the disk. Innermost in the disk (starting at $< 1000 R_s$) is a thick hot disk where radiation pressure exceeds the inward forces and the surface expands locally. This gives a concave surface to the mid disk that can be radiated by the thick hot disk. Beyond $100,000 R_s$, the disk breaks into clouds.

The black hole's Schwarzschild or Kerr⁶¹⁶ radius is tiny compared to that of the torus. For a SMBH mass of $10^9 M_\odot$, the Schwarzschild radius is 3×10^{12} m or about 0.1 light day or (1/2 this for Kerr equatorial radius for maximal rotation—IMA2 p. 641).

The radius of the torus is hard to pin down. There are few images which resolve the core of an AGN. The putative torus may be perhaps 200 to 400 light years in radius (400 from 1/2 of diameter in HST image⁶¹⁷ of NGC4261, 228 from IMA2 p. 1121, the latter giving a ratio of torus radius to Schwarzschild radius of “nearly 8 orders of magnitude”).

The HST image of nearby galaxy NGC 4261 shows “a 800 light-year-wide [400 ly radius] spiral-shaped disk of dust fueling [the SMBH]”,⁶¹⁸ but the text does not use the word torus. IMA2 p. 1121 however describes this same image as the core of an elliptical radio galaxy and optical LINER, showing a bright nucleus surrounded by a large, obscuring *torus* whose radius measures about 70 pc (228 ly). The SMBH is estimated to have a mass of $5 \times 10^8 M_\odot$.⁶¹⁹

⁶¹⁵ http://www.auger.org/news/PRagn/about_AGN.html

⁶¹⁶ http://en.wikipedia.org/wiki/Kerr_metric

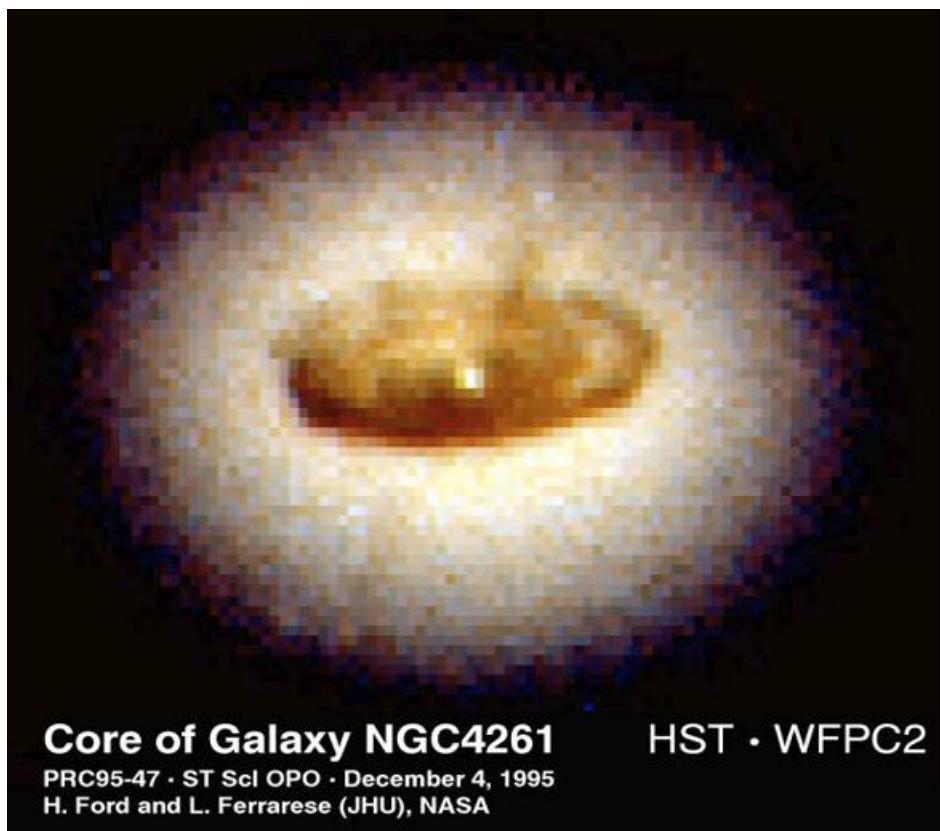
⁶¹⁷ http://www.mssl.ucl.ac.uk/www_astro/agn/agn_quasartour.html

⁶¹⁸ <http://hubblesite.org/newscenter/archive/releases/1995/47/text/>

⁶¹⁹ NGC 4261:

- <http://blackholes.stardate.org/directory/factsheet.php?p=NGC-4261>
- <http://www.stsci.edu/stsci/meetings/shst2/ferraresel.html>

For a maximally rotating $10^7 M_{\odot}$ SMBH, the estimated $R_{\text{Kerr}} = R_s/2 = 1.61 \times 10^{-6}$ lightyears. The estimated ratio of Torus radius (228 lightyears) to R_{Kerr} of the SMBH is roughly 1.5×10^8 , similar to the estimate in *IMA2*.



HST image of core at center of NGC4261, an elliptical radio galaxy and LINER, 1995 including an “800 light-year-wide [400 ly radius] spiral-shaped disk of dust”.

Mechanism for Producing a Relativistic Outflow of Charged Particles

The rotating conducting accretion disk can generate a magnetic field. The varying magnetic field may generate a large electric field near the surface of the disk, which causes acceleration of charged particles to relativistic speeds.

“Relativistic jets may provide evidence for the reality of frame-dragging. Gravitomagnetic forces produced by the Lense-Thirring effect (**frame dragging**) within the ergosphere of rotating black holes combined with the energy extraction mechanism by Penrose have been used to explain the observed properties of relativistic jets.”⁶²⁰

The relativistic outflow may also be produced by the **Blandford-Znajek effect**. With it, the accretion disk can tap the rotational energy of the BH itself. “This is the most popular theory for the extraction of energy from the central black hole. The magnetic fields around the accretion disk are dragged by the spin of the black hole. The relativistic material in the jets is possibly launched by the tightening of the field lines.”⁶²¹ The BH rotating in a magnetic field (of perhaps 1 T) can be thought of as an electrical conductor and generates an electrical potential between its poles and equator across a resistance of about 30Ω . A current flows, the BH spin is reduced, and power is generated that is imparted to the outflowing particles (paired electrons and positrons) and to EM radiation. Up to 9.2% of the rest energy of a maximally rotating BH can be extracted in this manner. The electrons spiraling about the magnetic field lines emit synchrotron radiation that

⁶²⁰ <http://en.wikipedia.org/wiki/Frame-dragging>

⁶²¹ http://en.wikipedia.org/wiki/Relativistic_jet

contributes to the continuous spectrum of AGNs. This synchrotron radiation spectrum follows a power law (*IMA2* p. 1115).

X-ray Generation

These are formed in several ways in addition to synchrotron radiation. The soft lower-energy X-rays may arise from the high-frequency end of the accretion disk spectrum. It is also possible that higher energy X-rays are produced by **inverse Compton scattering** of photons off the relativistic electrons⁶²²—this mechanism can produce very high energy photons, having final energy (from head-on collisions) of up to $4\gamma^2\hbar\omega_0$. Thermal bremsstrahlung scattering can also produce X-ray energy emissions.

Broad Line and Narrow Line Emission

These are the result of photoionization by the continuum radiation. The broad lines arise from permitted atomic transitions. The narrow lines may also involve forbidden transitions. Their variabilities differ, and these probably arise from different regions under different conditions.

The **broad line region** is close to the center of the disk. The broad-line emissions can respond rapidly to changes in the continuum radiation in as short as 1 week... They are clumpy, and contain partially ionized gas of about 10^{15} m^{-3} to 10^{16} m^{-3} electron number density. The temperature is about 10^4 K .

The narrow-line regions have much lower electron density, about 10^{10} m^{-3} , but have large overall mass and a temperature of about 10^4 K [same as broad-line region], and for these reasons can exhibit forbidden as well as allowed transitions. (*IMA2* p. 1116-1121)

Supermassive Black Hole Masses in AGNs versus Luminosities

These can be estimated using emissions from broad-line regions. Assuming a width and a corresponding rotational velocity of the lines (e.g., 5000 km s^{-1})⁶²³ and assuming a radius $r = 10^{15} \text{ m}$, the black hole mass may be estimated as

$$M_{\text{BH}} = rv^2/G = 1.9 \times 10^8 R_{\odot}.$$

Another technique uses **reverberation mapping**, in which the lag time (time delay τ) between changes in brightness of the continuum lines vs. the emission lines is evaluated. This technique may be most useful at high z and has the advantage of not requiring spatially resolving the central region.

The masses of SMBHs by reverberation mapping have been estimated at $0.2 \times 10^7 M_{\odot}$ to about $10^9 M_{\odot}$.

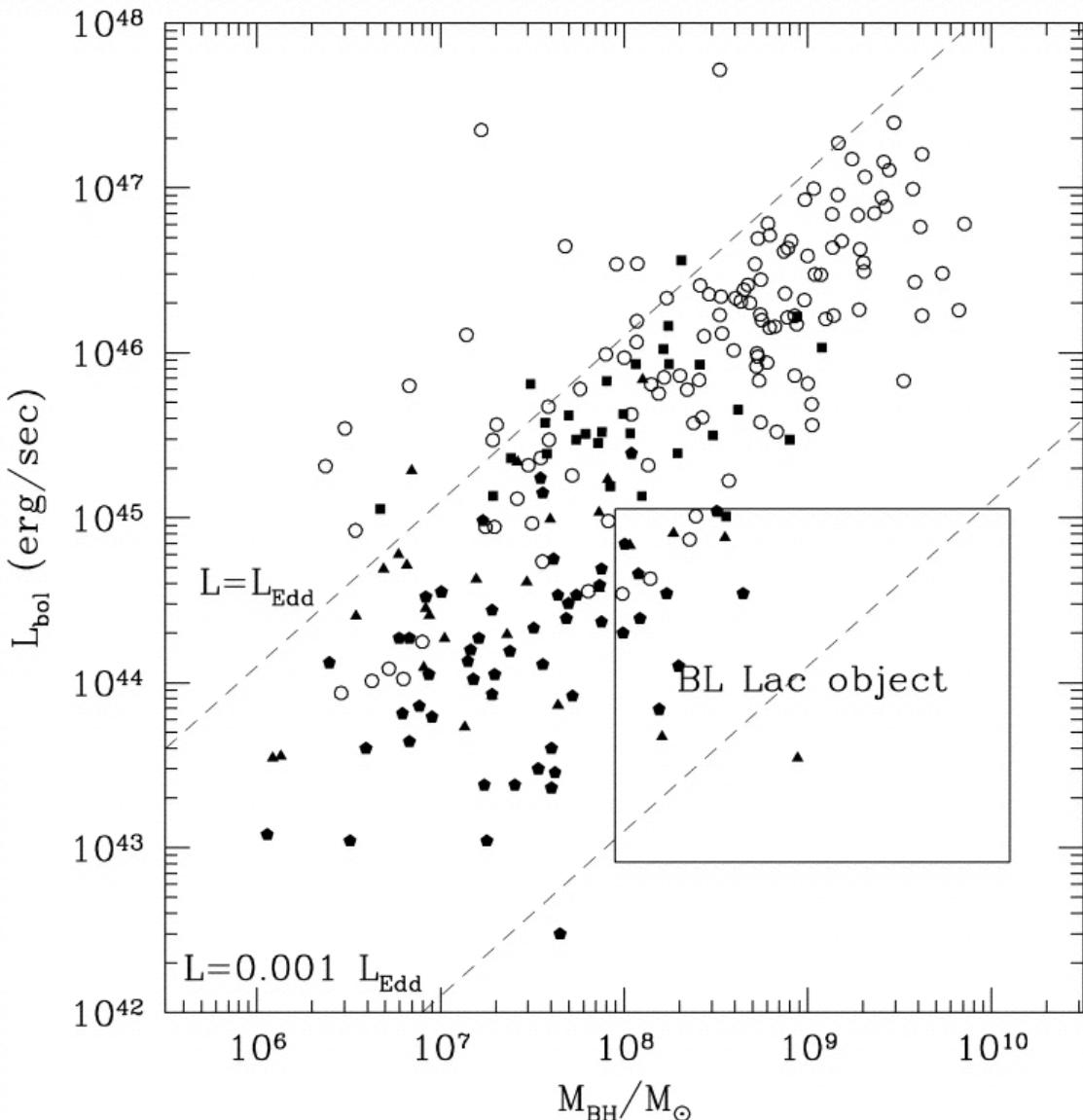
The bolometric luminosities of SMBHs in 234 AGNs have been shown in the following graph⁶²⁴ to only roughly correlate with their masses (see also somewhat different graph at *IMA2* p. 1118, which used by reverberation mapping).

⁶²² <http://www.cv.nrao.edu/course/astr534/InverseCompton.html>

⁶²³ “Doppler widths in the range ~ 1000 to $25,000 \text{ km s}^{-1}$ per http://www.astronomy.ohio-state.edu/~peterson/Docs/chapter_3.pdf

⁶²⁴ Woo Jong-Hak, C. Megan Urry, “Active Galactic Nucleus Black Hole Masses and Bolometric Luminosities”, *ApJ*, 579:530-544, 2002 November 10,

http://iopscience.iop.org/0004-637X/579/2/530/pdf/0004-637X_579_2_530.pdf



"Bolometric luminosity vs. black hole mass for 234 AGNs [using masses that were derived from optical luminosity and broad-line velocity]. There is little if any correlation... For a given black hole mass, there is a large range of bolometric luminosities, spanning 3 or more orders of magnitude. The Eddington limit [upper dashed line $L=L_{\text{Edd}}$] defines an approximate upper limit to the luminosity, but the absence of objects from the lower right of the diagram (low-luminosity, high-mass AGNs) is a selection effect. For example, this part of the diagram would be occupied by BL Lac objects and low-luminosity radio galaxies..." The rectangle delineates where BL Lacs are expected, but the authors note that for BL Lacs, "we do not have good estimates of bolometric luminosity".
 (Woo and Urry, 2002)⁶²⁵

Summary of the Unified Model, With Emphasis on the Observer's Angle of View

According to the unified model, the central engine of an AGN is a supermassive black hole with a surrounding accretion disk. It is powered by conversion of gravitational potential energy to rotational kinetic energy, thermal energy, and synchrotron radiation. The structure of the disk (e.g., how puffed up it is) depends on the ratio of accretion disk luminosity and the Eddington luminosity limit (which is mass-dependent). The accretion rates typically are 1 to $10 M_{\odot} \text{ yr}^{-1}$. The black hole's Schwarzschild radius or Kerr⁶²⁶ radius is tiny compared to that of the torus (see details above). For a SMBH mass of $10^9 M_{\odot}$, the Schwarzschild radius is

⁶²⁵ ibid.

⁶²⁶ http://en.wikipedia.org/wiki/Kerr_metric

about 0.1 light day (1/2 this as Kerr equatorial radius for maximal rotation). The torus radius may be perhaps 200 to 400 light years. The jets reach out some 15 kpc from the nucleus.

The following appearances based on view perspective are postulated (see diagrams):

- Seyfert 2, NLRG: seeing the disk from the side is blocked by the obscuring torus so can't see the AGN or its radiation directly, and can't see Broad-line Region contained inside the torus. The Narrow Line Region is positioned further out, so its emissions are not obscured by the torus.
- Seyfert 1, BLRG, QSO: viewing from about halfway between torus and central axis, sometimes on the no/weak jet side, can see the AGN partially, greater luminosity. Can see Broad-line region + Narrow Line Region. Radio is weak w Seyfert 1 & QSO, strong with BLRG.
- Blazars (BL Lac, OVV), Radio Loud Quasars QSR: viewing close to (for QSRs) or directly on (for Blazars) the strongest jet axis. Strong radio (esp. for Blazars), strong continuum, broad + narrow emission lines.
- Radio Quiet Quasars QSO: Can also arise from viewing from the no/weak jet side.

AGN Evolution

Merger with other galaxies and companions help to provide fuel to the accretion disk of an Active Galactic Nucleus (*IMA2* p. 1129). As an AGN runs out of fuel, the central luminosity may decline before the radio lobes exhibit decreasing emission. AGNs may undergo periods of quiescence punctuated by reactivation from mergers, etc. The lifetimes of AGNs are not known.

Radio Lobes and Jets

Radio-loud galaxies also tend to be strong emitters in X-ray.

Generation of Jets and Radio Lobes

The jets consist of charged particles that have been accelerated away from the nucleus to relativistic speeds in opposite directions. *IMA2* states that the jets may be electrons and positrons or electrons and ions. The magnetic field of the disk becomes *frozen into* the plasma of the jets of outgoing charged particles. The remarkable collimation of some jets suggests that the process takes place close to the central engine powering the jet. The collimation may arise from the hot accretion disk around the SMBH. Material will pile up on the inner “centrifugal barrier”, leaving an empty nozzle more centrally through which the jet may flow. Poorly understood magnetohydrodynamic MHD effects may also play a role.

A jet encounters resistance in the interstellar medium within the galaxy and in the intergalactic medium beyond. A shock front develops and the jets slows—some of the energy “splashes back” and also becomes disordered. The highly turbulent nature of this disorder is quite difficult to model.

A single lobe (such as Cygnus A) may hold 10^{53} to 10^{54} joules as kinetic and magnetic energy (the latter given by $\mu_M = B^2/2\mu_0$ in units of $J \text{ m}^{-3}$). Given this store of energy and the observed luminosity rates, it is possible that a radio lobe might have a lifetime of more than 10^8 years. *IMA2* p. 1125 estimates the magnetic field in a radio lobe at 41 nT (range about 1 nT to 100 nT).

The jet radiation emissions exhibit a power law spectrum and strong linear polarization, consistent with origin from synchrotron radiation. Particles may be accelerated by magnetic squeezing by shock waves, and also radiation pressure.

Relativistic Beaming

Jets often exhibit **relativistic beaming** or **headlight effect**, another consequence of special relativity, specifically the Lorentz transformation of velocities. Rays of light emitted by a moving object appear to a relatively nonmoving observer as if they have arisen concentrated into a forward angle more in the direction of the object's motion. Thus light emitted by a moving object appears to be compressed or concentrated into a cone about its direction of motion, an effect called relativistic beaming, and this can intensify the brightness

of the beam. This concentration of light in the object's forward direction is also referred to as the headlight effect. (IMA2 p. 101-2) This contributes to the appearance of one-sided beams when opposing beams are more likely.

Apparently Superluminal Jets

Some jets appear to have **superluminal** velocity exceeding c . For example, 3C 273 exhibits knots (blobs) with an angular velocity of $\mu = 0.0008'' \text{ yr}^{-1}$. The distance to 3C 273 is $440 h^{-1} \text{ Mpc}$. Then assuming $h = h_{\text{WMAP}}$, the apparent transverse velocity = $v_{\text{app}} = \mu d = 7.85c$. The explanation involves special circumstances in which the blob has been ejected at relativistic speed so that a narrow angle φ is formed between the jet direction and the observer's line of sight. "This phenomenon is caused because the jets are travelling very near the speed of light and at a very small angle towards the observer. Because at every point of their path the high-velocity jets are emitting light, the light they emit does not approach the observer much more quickly than the jet itself. To be more clear, the jet is essentially 'chasing' the light it emits. This causes the light emitted over hundreds of years of travel to not have hundreds of lightyears of distance between it, the light thus arrives at the observer over a much smaller time period (ten or twenty years) giving the illusion of faster than light travel."⁶²⁷ The maximal apparent transverse velocity depends on the angle and the actual velocity of the jet. The formula is

$$\beta_T^{\max} = \beta\gamma$$

where $\beta_T = v_T/c$ (the apparent transverse velocity divided by c), and γ = Lorentz factor. If $\gamma \gg 1$ and the jet velocity v is close to c , then $\beta_T^{\max} > 1$.

Using Quasars to Probe the Universe

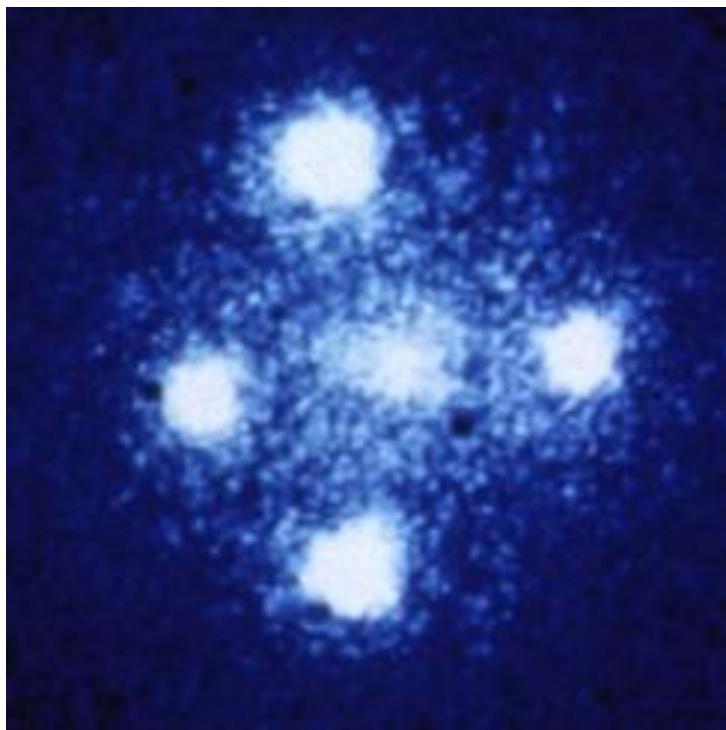
Gravitational Lensing

QSRs are among the most distant objects known. Fritz Zwicky (1898-1974) suggested that galaxies could cause gravitational lensing. Many examples have been found. The image of a QSR lensed by a galaxy that lies exactly on the line of sight will be an Einstein Ring.⁶²⁸ If the lensing object is off axis, the ring will appear as a partial ring or arc.

If the lens is extended and elliptical in shape, 3 or 5 images can be formed, for instance, the **Einstein cross** (G2237+0305 = QSO 2237+0305).⁶²⁹

Apparently there are at least two other Einstein crosses, but I have not yet located information on them.⁶³⁰

Image to right is of the Einstein Cross Gravitational Lens. This is a gravitationally lensed quasar that sits directly behind galaxy ZW 2237+030. Four images of the same distant quasar appear around the foreground galaxy (the latter probably obscures a 5th quasar image) due to strong gravitational lensing. The angular separation between the upper



⁶²⁷ http://en.wikipedia.org/wiki/Superluminal_jet

⁶²⁸ http://en.wikipedia.org/wiki/Einstein_ring

⁶²⁹ http://en.wikipedia.org/wiki/Einstein_Cross

⁶³⁰ <http://scienceworld.wolfram.com/physics/EinsteinCross.html>

and lower images is 1.6 arcseconds. (HST using ESA Faint Object Camera)⁶³¹

Light from distant galaxies or quasars passing through galaxy clusters (such as Abell 1689 or Abell 2218) may be distorted into a series of arcs and other deformed shapes.⁶³² The **Einstein Circle or Einstein Ring** is a special case of gravitational lensing, caused by the exact alignment of the source, lens and observer (which almost never occurs). This results in a symmetric ring around the lens. The angle of the ring in radians is given by the **Einstein Radius θ_E :**⁶³³

$$\theta_E = \sqrt{\left(\frac{4GM}{c^2}\right)\left(\frac{D_{LS}}{D_L D_S}\right)}$$

where D_{LS} = Angular Diameter Distance from Lens to Source, D_L = Angular Diameter Distance from lens to observer, and D_S = Angular Diameter Distance from source to observer. Angular Diameter Distance is a relativity-related measure. Note that for nearer objects, $D_{LS} = D_S - D_L$, but at cosmological distances, $D_{LS} \approx D_S - D_L$ due to relativity effects. If the source is off-axis, multiple arcs or images may be seen rather than a ring.

The multiple lensed images of a quasar may vary in relative intensity with time. In addition, the images are taking different paths in space so that their times of arrival for the same source time may differ by more than a year, further complicating things when the source is variable.

The differing delay among multiple images has been used as an independent estimate of the Hubble constant H_0 (69 vs. 71 km s⁻¹ Mpc⁻¹ by WMAP), using for instance the Twin Quasar,⁶³⁴ for which the distant quasar is at $z = 1.4$. The two quasar images are separated by 6 arcsec, and there is a 417 day time lag between the two images.⁶³⁵

Gravitational Microlensing is proving useful, for example in the Microlensing Planet Search Project (MPS): "Planetary systems in the foreground of the Galactic Bulge form a class of (planetary) binary lens systems that can be detected through photometric microlensing experiments. An Earth mass planet orbiting a lensing star can perturb the would-be single lens microlensing light curve due to the lensing star, briefly but spectacularly. Gravitational microlensing is the only known ground-based method to probe earth mass planets orbiting around main sequence stars. (Earth-mass objects orbiting a neutron star were discovered in 1992 by Wolszczan and Frail.)"⁶³⁶

David Darling states: "When aligned with a more distant star, a microlens deflects, distorts, and brightens the image of the background source in a way that is sensitive to the strength and geometry of the lensing star's gravitational field. If the lensing star is orbited by a planet then the result can be a microlensing anomaly known as a **binary lens**, one of the characteristics of which is the appearance of so-called **caustic curves** in the plane of the source. As the source sweeps across the caustic curves, due to the relative motion of source and lens, sharp peaks appear in the light curve. Planets located at a distance between about 0.5 and 1.5 times the Einstein radius of their host star are said to be in the "lensing zone," since they are most likely to produce, with the help of the host, a sizable caustic structure. This distance is typically a few astronomical units, or approximately the distance of Jupiter from the Sun. Calculations (for example, by Bohdan Paczynski) suggest that microlensing anomalies due to Jupiter-mass planets in Jupiter-like orbits will last about 1 to 3 days and be detectable in 10 to 30 percent of cases where such planets are in orbit around a lensing star that is located roughly halfway between us and the center of the Galaxy. Smaller mass planets will create smaller caustic structures which (1) reduces the chance of the background source passing directly behind the anomaly and (2), in cases where the source does pass behind the anomaly, results in a shorter and less sharp increase in brightness. The probability of detection and the duration of the lensing peak decrease as the square root of the planet's mass. So, an Earth-mass planet (320 times less massive than Jupiter), orbiting at about the radius of its host's Einstein ring, will give rise to a caustic structure 18 times

⁶³¹ <http://hubblesite.org/newscenter/archive/releases/1990/20/image/a/>

⁶³² http://en.wikipedia.org/wiki/Gravitational_lens

⁶³³ http://en.wikipedia.org/wiki/Einstein_radius

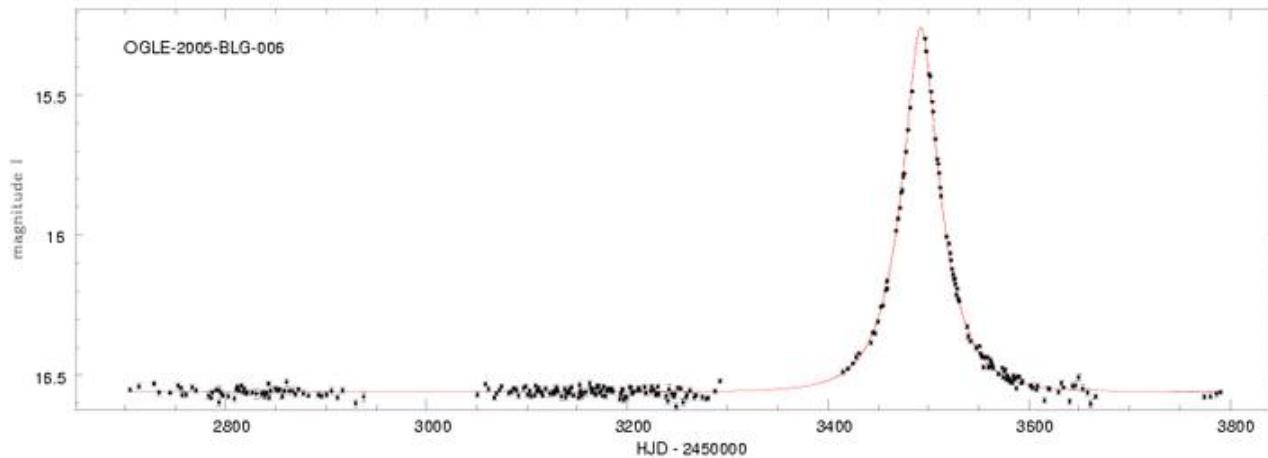
⁶³⁴ Twin Quasar: (Q0957+561) or (QSO 0957+561 A/B) or (SBS 0957+561)

⁶³⁵ http://en.wikipedia.org/wiki/Twin_Quasar

⁶³⁶ <http://bustard.phys.nd.edu/MPS/>

smaller than that of a Jupiter-mass planet in a similar orbit, and any light curve variations caused by it will last for only a few hours rather than a few days.”⁶³⁷

To emphasize, microlensing is manifested not by image splitting, distortion, or angular magnification, but by a smooth brightening (loosely spoken of as “magnification”) and dimming of the lensed source relative to its unlensed state. (Image that follows taken from *here*⁶³⁸):



The search for exoplanets using this state-of-the art technique and other research techniques warrants much more attention when I find the time.

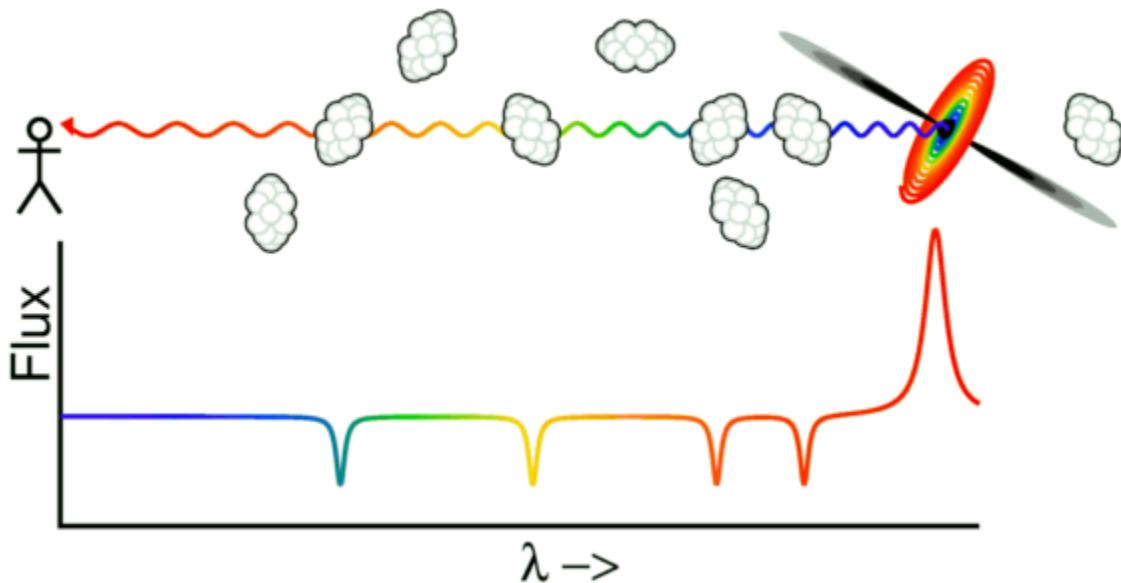
The Lyman-Alpha Forest

The resting Lyman-alpha line⁶³⁹ is at 1215.668 Angstroms or 121.5668 nm, thus in the vacuum UV. High redshift quasars usually emit a prominent Lyman- α emission line, which is strongly redshifted by the time it reaches us (*IMA2* shows a Lyman- α emission line for QSO 1215+333 at about 440 nm). When broad-spectrum light from the distant quasar passes through multiple H I clouds of neutral hydrogen, absorption transitions occur at the Lyman- α wavelength at each location. These create absorption lines that are then redshifted as the light propagates toward the observer, with greater redshift in absorption lines occurring for clouds closer to the source. Each individual cloud leaves its own spike-like Lyman- α absorption line at a different position in the observed spectrum, and the numerous and variably redshifted Lyman- α absorption lines taken as a group are termed the *Lyman-Alpha Forest*.

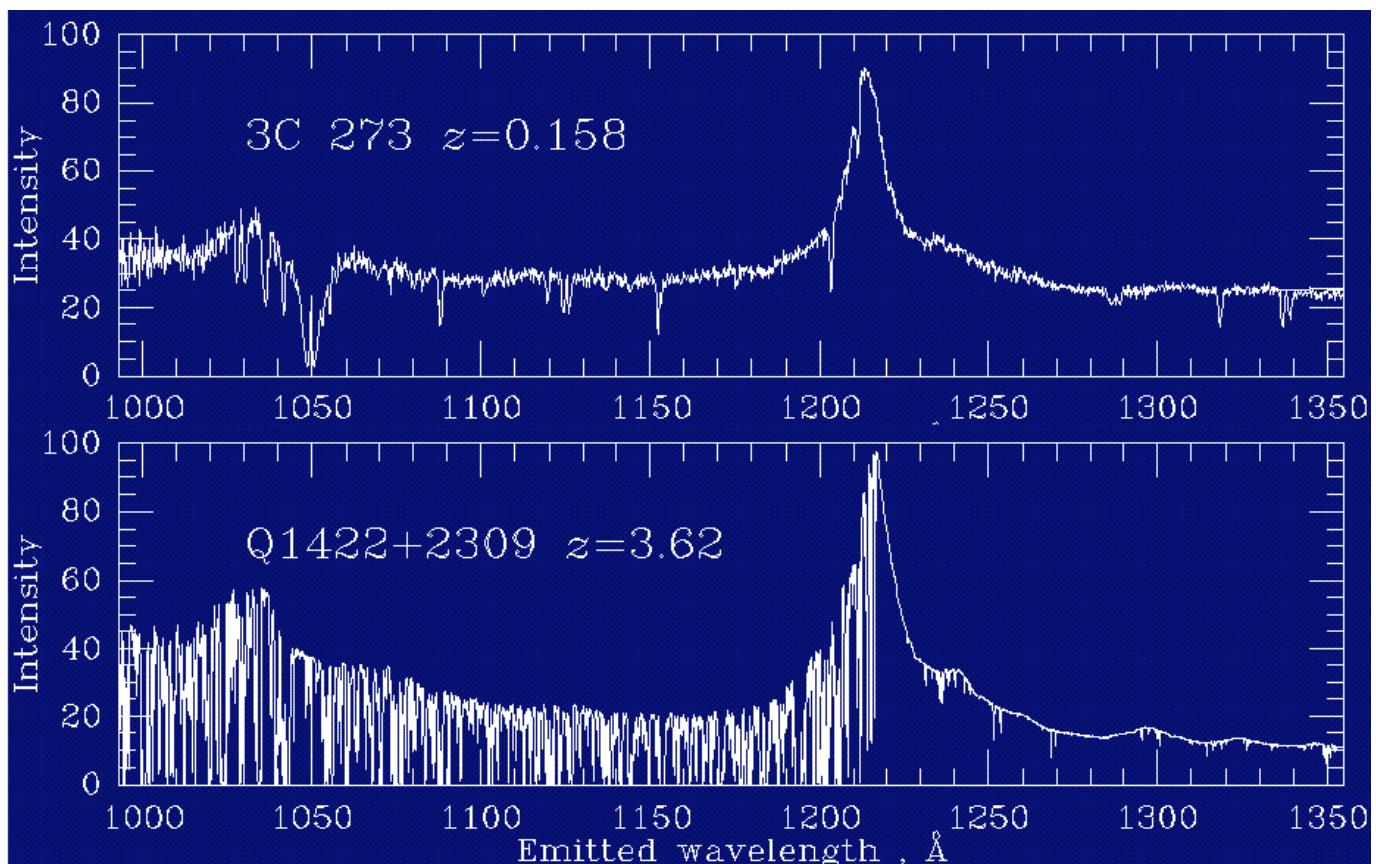
⁶³⁷ <http://www.daviddarling.info/encyclopedia/M/microlensing.html>

⁶³⁸ http://en.wikipedia.org/wiki/Gravitational_microlensing

⁶³⁹ http://en.wikipedia.org/wiki/Lyman-alpha_line



Cartoon showing single maximally redshifted Lyman- α emission peak (red on the right), and a few of the often many Lyman-Alpha Forest absorption lines (various colors) arising from clouds positioned at varying redshift distances between source quasar and observer. Absorption lines closer to the observer are less redshifted and depicted as blue or green here.⁶⁴⁰



Spectra of nearby quasar 3C 273 ($z = 0.158$) and of a distant quasar Q1422+2309 ($z = 3.62$).

The spectra are scaled to a common scale in emitted wavelength so that the Lyman-alpha emission peak is shown at the same wavelength, $\sim 1216 \text{ \AA}$, in both cases.

The Lyman-alpha forest of hundreds of absorption lines for the more distant quasar is apparent at the lower left. It has been a long journey!⁶⁴¹

⁶⁴⁰ <http://www.astro.ucla.edu/~wright/Lyman-alpha-forest.html>, Edward Wright 2004.

⁶⁴¹ modified slightly by MCM from <http://www.astr.ua.edu/keel/agn/forest.gif>, by Bill Keel

Ionized Metal Absorption Lines in Quasars

The narrow absorption lines of ionized metals seen in ground-based quasar spectra vary by z . For $z < 1.5$, the Mg II lines dominate in ground based spectra (and are accompanied by Si II, C II, Fe II, and ASI II lines. For $1.2 < z < 3.5$, however, C IV, Si IV, N V, and O IV lines are commonly seen. The latter may be produced by strong ionization of clouds by OB stars in young galaxies. These clouds also have lower abundances of heavy elements, consistent with their arising in young galaxies.

Intergalactic Clouds

The comoving space density of intergalactic clouds appears to have been substantially greater in the past. They are not grouped in clusters but appear to be distributed randomly.

Interesting simulations by Renyue Cen of spatial density of clouds of H I may be seen *here*.⁶⁴² (The simulations also include many other parameters, apparently in a comoving cubical volume 25-Mpc on a side, although the details are beyond my skill level) Look for instance at H I density at $z = 6$ down to $z = 0$, a sequence clearly showing the decline of density with lower z (younger age).

Incidentally, Cen also presents the following elegant simulation⁶⁴³ of galaxy formation, set to a lovely O Mio Babbino Caro soundtrack.

⁶⁴² <http://www.astro.princeton.edu/~cen/PROJECTS/p2/p2.html>

⁶⁴³ http://www.astro.princeton.edu/~cen/AMR_GalaxyFormation-2.mp4

Cosmology (Chapter 29, omitted)

The Early Universe (Chapter 30, omitted)

(omitted, insufficient time, maybe next year)