DENVER MUSEUM OF NATURE & SCIENCE

Technical Report 2005–10

Ka Chun Yu

Lecture Notes for Introduction to Astronomy

Denver – 2005

DMNS Technical Report 2005-10, 216 pp.

Lecture Notes for Introduction to Astronomy

September–November, 2004

by

Dr. Ka Chun Yu

Curator of Space Science

Denver Museum of Nature & Science

Contents

1 The Night Sky **1**
- 1.1 Overview . . . 1
- 1.2 The Celestial Sphere and Daily Motions . . . 2
  - 1.2.1 The Autumn Sky . . . 4
  - 1.2.2 The Winter Sky . . . 7
  - 1.2.3 The Spring Sky . . . 9
  - 1.2.4 The Summer Sky . . . 11
- 1.3 The Stars from Different Latitudes . . . 14
- 1.4 Star Motions Through the Year . . . 15
- 1.5 The Seasons . . . 18
- 1.6 The Phases of the Moon . . . 20

2 The Terrestrial Planets **23**
- 2.1 The Ancient View of the Planets . . . 23
- 2.2 The Copernican Revolution . . . 25
- 2.3 Origin of the Solar System . . . 30
  - 2.3.1 Cratering . . . 34
- 2.4 The Structure of the Earth . . . 35
  - 2.4.1 Crust, Mantle, and Core . . . 35
  - 2.4.2 Plate Tectonics . . . 36
- 2.5 Comparative Planetology . . . 42
  - 2.5.1 The Moon . . . 42
  - 2.5.2 Mercury . . . 44
  - 2.5.3 Venus . . . 44
  - 2.5.4 Mars . . . 46
- 2.6 Terrestrial Atmospheres . . . 46
  - 2.6.1 Venus . . . 46
  - 2.6.2 Earth . . . 54
  - 2.6.3 Mars . . . 54

3 The Outer Solar System **57**
- 3.1 Jupiter . . . 58

3.1.1 The Galilean Moons . . . 64
3.1.2 Tidal Heating . . . 65
3.2 Saturn . . . 68
3.3 Uranus and Neptune . . . 72
3.4 Pluto and the Kuiper Belt . . . 76
3.5 Comets . . . 77

4 The Sun 83
4.1 Nuclear Fusion at the Sun's Core . . . 89
4.2 The Radiative and Convective Zones . . . 90
4.3 The Solar Atmosphere . . . 93

5 Stars and the Milky Way 97
5.1 The Life Cycle of Stars . . . 97
5.2 Stars: Building Blocks of Galaxies . . . 99
5.3 Star Formation . . . 102
5.4 A Star's Main Sequence Life . . . 114
5.5 Post Main Sequence Low Mass Stars . . . 116
5.6 Stellar End-Life: High Mass Stars . . . 119
5.7 Structure of the Milky Way . . . 125
5.8 The Evolution of the Milky Way . . . 127

6 The Universe and the Big Bang 133
6.1 The Extra-Galactic Zoo . . . 133
6.1.1 Normal Galaxies . . . 133
6.1.2 Active Galaxies . . . 142
6.1.3 Galaxy Clusters and Superclusters . . . 146
6.1.4 Galaxy Cluster Mass . . . 156
6.1.5 Dark Matter . . . 157
6.2 The Big Bang . . . 161
6.2.1 The Expanding Universe . . . 161
6.2.2 Cosmic Element Abundances . . . 162
6.2.3 The Cosmic Microwave Background . . . 163
6.3 Inflation . . . 167
6.3.1 Problems With the Big Bang . . . 167
6.3.2 Problems Solved? . . . 169
6.4 The Contents and Fate of the Universe . . . 171

A Notation 177
A.1 Scientific Notation . . . 177
A.1.1 "Orders of Magnitude" . . . 178
A.2 Scientific Units . . . 179
A.3 Astronomical Units . . . 180

B The Scientific Process **183**

C Electromagnetic Radiation **189**
C.1 Electromagnetic Radiation 189
C.2 Spectral Lines 191
C.3 Blackbody Radiation 195

Further Reading **199**

List of Figures

1.1 Finding Polaris 3
1.2 Cassiopeia 5
1.3 Cassiopeia, Cepheus Perseus, Andromeda, and Pegasus 6
1.4 The Winter Sky 8
1.5 The Winter Sky: Leo and Cancer 9
1.6 Arcing to Arcturus 10
1.7 Scorpio 11
1.8 The Summer Triangle 12
1.9 Sagittarius 13
1.10 The Celestial Equator and Celestial Poles 15
1.11 September Night Sky 16
1.12 September Day Sky 17
1.13 The Sun at Different Times of the Year 19
1.14 The Seasons on Earth 21
1.15 Phases of the Moon 22

2.1 Aristotle and Ptolemy 23
2.2 Ptolemy's Geocentric Universe 24
2.3 Epicycles and Deferents 25
2.4 Nicolaus Copernicus 26
2.5 *De Revolutionibus* 27
2.6 Johannes Kepler and Galileo Galilei 28
2.7 Ellipses 28
2.8 Grain and Planetesimal Growth 32
2.9 Accretion in the Solar System 33
2.10 Condensation Sequence 34
2.11 Cutaway of the Earth 35
2.12 Convection in the Mantle System 37
2.13 Plate Motions 38
2.14 Hawaiian Islands 39
2.15 Global Tectonic Plates 40
2.16 The Break Up of a Super Continent 41
2.17 Formation of the Moon 43

2.18 The Surface of Mercury 45
2.19 Magellan Images of Venus 47
2.18 Magellan Images of Venus (Cont.) 48
2.18 Magellan Images of Venus (Cont.) 49
2.18 Magellan Images of Venus (Cont.) 50
2.18 Magellan Images of Venus (Cont.) 51
2.20 Valle Marineris 52
2.21 The Tharsis Bulge 53

3.1 Jupiter 59
3.2 Io 60
3.3 Europa 61
3.4 Ganymede 62
3.5 Callisto 63
3.6 Volcanoes on Io 66
3.7 Ice Rafts on Europa 67
3.8 Rings Close-Up I 70
3.9 Ring Close-Up II 71
3.10 Moons of Saturn 73
3.11 Uranus 74
3.12 Neptune 75
3.13 Pluto and Charon 76
3.14 Comparison of Outer Solar System Bodies 78
3.15 Comet C/2001 Q4 80
3.16 Nucleus of Halley's Comet 81
3.17 The Kuiper Belt and Oort Cloud 82

4.1 The Current Sun: Visible Wavelengths 84
4.2 The Current Sun: X-Ray Wavelengths 85
4.3 Fraunhofer Lines 87
4.4 Structure of the Sun 88
4.5 Convection Cells Inside the Sun 91
4.6 Granules in the Photosphere 92
4.7 The Solar Chromosphere 94
4.8 Coronal Mass Ejection 95

5.1 M51 98
5.2 Barnard 68 103
5.3 The Perseus Molecular Cloud Complex 104
5.4 Star Formation: The Giant Molecular Cloud 106
5.5 Star Formation: Embedded Clusters 107
5.6 Star Formation: Emerging Clusters 108
5.7 Star Formation: Low Mass Stars with Outflows 109

5.8 Star Formation: The End of the GMC . . . 110
5.9 BHR 71 . . . 111
5.10 The Orion Star Forming Region . . . 112
5.11 RCW 49 . . . 113
5.12 Hydrostatic Equilibrium . . . 115
5.13 A Main Sequence Star . . . 115
5.14 Heliuim Shell Burning . . . 117
5.15 Comparison Between a Red Giant and Its Main Sequence Precursor . . . 118
5.16 A Carbon Core Star . . . 118
5.17 Planetary Nebulae . . . 120
5.18 The Onion Star . . . 121
5.19 Crab Nebula . . . 123
5.20 Heart of the Crab Nebula . . . 124
5.21 Black Hole . . . 125
5.22 The Milky Way . . . 126
5.23 The Galactic Fountain . . . 128
5.24 Growth of the Milky Way . . . 129
5.25 Sagittarius Dwarf Galaxy . . . 131

6.1 Hubble Deep Field . . . 134
6.2 Elliptical Galaxies . . . 135
6.3 Unbarred Spiral Galaxies . . . 137
6.4 Barred Spiral Galaxies . . . 138
6.5 Lenticular Galaxies . . . 139
6.6 The Magellanic Clouds . . . 140
6.7 More Irregular Galaxies . . . 141
6.8 Hubble's "Tuning Fork" Classification . . . 142
6.9 Comparison of Active and Normal Galaxies . . . 143
6.10 3C273 . . . 144
6.11 Quasar Host Galaxies . . . 145
6.12 Model of an AGN . . . 146
6.13 Satellite Galaxies of the Milky Way . . . 147
6.14 The Local Group . . . 148
6.15 The Virgo Supercluster . . . 149
6.16 The Virgo Cluster . . . 150
6.17 The Coma Cluster . . . 151
6.18 Neighboring Superclusters . . . 152
6.19 2-Degree Field Galaxy Survey . . . 153
6.20 The APM Survey of Galaxies . . . 154
6.21 The Observable Universe . . . 155
6.22 Hydra A Galaxy Cluster . . . 157
6.23 Gravitational Lensing by Abell 2218 . . . 158

6.24 Rotation Curve of the Milky Way 159
6.25 Hubble's Law 162
6.26 The Expanding Wavelength of a Photon 164
6.27 Seeing to the Edge of the Universe 165
6.28 COBE Spectrum 166
6.29 Curvature of Three Possible Universes 168
6.30 Expansion from Inflation 170
6.31 Inflation Solves the Flatness Problem 170
6.32 The Accelerating Universe 173

C.1 The Bohr Model of the Atom 192
C.2 Electron Excitation 194
C.3 Blackbody Spectra at Different Temperatures 196
C.4 A Variety of Blackbody Emitters 197

List of Tables

2.1 Planet Distances 29
2.2 Planet Periods 30
2.3 Composition of Terrestrial Atmospheres 55

3.1 Pluto/Charon and the Kuiper Belt Objects 77

4.1 Solar Composition 86

5.1 Stellar Properties 100

6.1 Baryonic Census 174

Lecture 1

The Night Sky

1.1 Overview

This course will take place in the planetarium during the next six weeks. For the first lecture, we will go over the night sky as seen from the Earth, and talk about the patterns in the sky that are visible, how they change over time, and how the change their appearance when we move around on the Earth.

The two weeks following, we will move out into the Solar System and start exploring our neighboring planets. First we will cover the four inner terrestrial planets: the Earth, Mars, Venus, and Mercury. We will talk about specifics of their surfaces and atmospheres and compare their different geologies and activity. The next week, we will move into the outer solar system where we will focus on the four gas giant planets, their icy moons, and the host of other rocky and icy bodies and debris in the outer Solar System. This latter category includes comets in the Oort Cloud as well as the Kuiper Belt objects.

We finish the Solar System with a Lecture on the Sun, and how it generates its energy. In an important aside, we will also discuss the physical processes by which light is emitted and absorbed by atoms. This is critical for understanding how we learn about the properties of distant objects if we can only sample their light to study.

In Lecture 4, we will move out of the Solar System and talk about the Milky Way Galaxy, and its main building block, the stars. We'll look at their life histories, the places they are born in, how they die, and how the life cycle of stars connects them with the interstellar medium. We will also discuss the major components of the Milky Way, as well as a theory about its origins.

In the last week, we are going to go on a whirlwind of a lecture on the whole rest of the universe, including looking at all the different types of galaxies and galaxy morphologies that we observe, a look at galaxy clusters, and clusters of galaxy clusters which turn out to be some of the largest structures in the visible universe. We will conclude with the Big Bang, and examine just some of the observational evidence that supports it. We will finish with some very recent results from cosmology that give rather startling revelations.

This is an introductory astronomy class, and as such, there are no prerequisites of any kind. I won't suppose that you have had any previous astronomical knowledge before. This class should be understandable even if you've never taken a science or math class before in your life (or perhaps it's been a decade or more since the last time that you have). Of course that will make my job a little more difficult, and I will try to make everything as understandable as possible to the average layperson. We will be covering an enormous amount of material, and it will be impossible to do any one topic justice. However in the end, I hope that you will have a deeper appreciation about the astronomical wonders around us, and have a very broad, albeit somewhat shallow, level of understanding of our universe.

1.2 The Celestial Sphere and Daily Motions

Imagine going back in time thousands of years, before there was the Internet, or cell phones, or cable or broadcast TV, or movies, or radio, or even electricity. Even just a hundred years ago, it was much easier to see the night sky. Lighting at night wasn't ubiquitous as it is now; although you could have campfires or torches, they didn't cast their light over a very wide area. All you had to do to see the night sky was to step away from these small pools of illumination and the sky would just be there, its vast expanse above and all around you.

It is obvious to us even as young children that the Sun moves in the sky. It rises in the east, transits across the sky, and sets in the west. Without the distractions of modern day city life, it would be obvious quickly to us all that the night sky moves as well, in a very regular and distinct way. Although some stars rose and made great big arcs in the sky, other stars hardly moved at all. These instead made small circular motions about a single point in the heavens. And for a star that was exactly at that point, it appeared to be stationary throughout the night.

It would appear that everything in the sky wheeled about this single point, as if the stars and the Sun were attached to the surface of a giant globe or dome that surrounds the Earth. This globe, or **celestial sphere** appears to spin about an axis. For nearly everyone in the Northern Hemisphere, one point of this axis is visible and is called the **north celestial pole**; and for people living in the Southern Hemisphere, they saw the opposite end of the axis of the spinning celestial sphere, and that is the **south celestial pole**.

Luckily for people in the Northern Hemisphere, there was a star that was almost exactly at the north celestial pole, and that is **Polaris**, the Pole Star. Thus throughout the night and throughout the year, Polaris will always be up in the night sky, and for most people living in the Northern Hemisphere, it was always located north of them. As a result, Polaris is also called the **North Star** and has been used for millenia as a navigator's tool in finding true North.

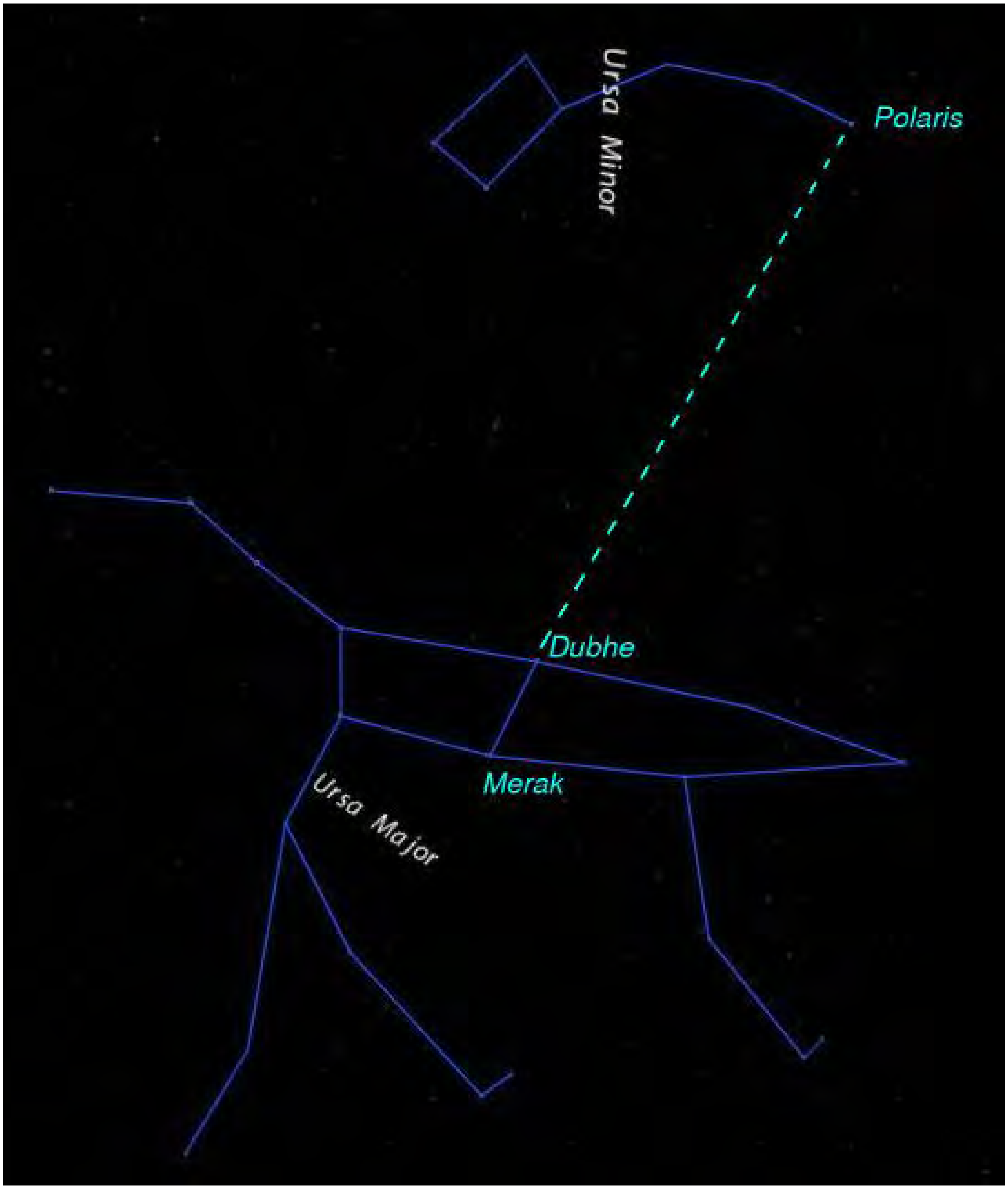


Figure 1.1: Finding Polaris

After locating the Big Dipper (or Ursa Major), draw a line between the two Pointer Stars, Dubhe and Merak, and follow it to the North Star, Polaris.

When we turn on the Cosmic Atlas calendar tool, we see that the time and date we have set for ourselves is sometime tonight[1]. Now as we move time forward quicker, the stars will rise in the east (to our left) and set in the west (to our right) faster. The stars further to the south (ahead of us) move in great big arcs as they rise and set; however the stars to the north, behind us, move in much smaller arcs. As you look closer to Polaris, the stars appear to move in circular motions with the circles centered about Polaris.

As we learn about the night sky, Polaris is a useful starting point since that is always visible every night throughout the year. However it is a somewhat faint star and it is hard to see. Luckily there is another set of stars that we can use to help us find Polaris. This is the **Big Dipper** which is also visible for most of the night for people living in the Northern Hemisphere. The Big Dipper is easy enough to make out as a constellation, and is bright enough that almost everyone can find it in the sky. When you go home tonight, see if you can find it (if it is visible and not covered by clouds).

After you locate the Big Dipper, look for the two bright stars that make up the far edge of the bowl in the Dipper. These two stars are the two brightest stars in the Big Dipper, **Dubhe** and **Merak**. They are also called the **Pointer Stars** because if you connect them with a line and follow it above the Big Dipper, you will hit Polaris. After you find Polaris, which sits at the end of the handle of the **Little Dipper**, you will be able to trace out the rest of its parent constellation.

1.2.1 The Autumn Sky

We will be able to proceed to find other constellations once we have located these first two. Notice that the Little Dipper appears upside down to us in the fall months. One American Indian myth explains that the autumn colors spill out of the Little Dipper when it is upside down, and that is what makes the leaves on trees turn colors in the fall.

Now let us follow the Pointer Stars past Polaris an equal distance away and you will find a set of stars shaped like a W. This is **Cassiopeia**, named after the Queen of the Ethiopians in Greek mythology. To the west of her is a set of fainter stars that are shaped like a house with a peaked roof. This is her husband, **Cepheus**. On the other side of Cassiopeia is the constellation of **Andromeda** which is made up of two trails of stars that form a big curved V in the sky. And just off to the side of the V of Andromeda is a fuzzy patch of sky, which is where the Great Galaxy in Andromeda is located. We will discuss galaxies at the very end of this course, but you should remember that if you are able to see the Andromeda Galaxy, you are looking at the farthest object that can be seen by the naked eye without the aid of binoculars

[1]The reason why the hour and date seems to be off is that it is currently using Greenwich mean time, or time as recorded at the Greenwich Observatory outside of London. Since London is 6 hours ahead of us, if you subtract 6 hours from the time shown, it is roughly the time now.

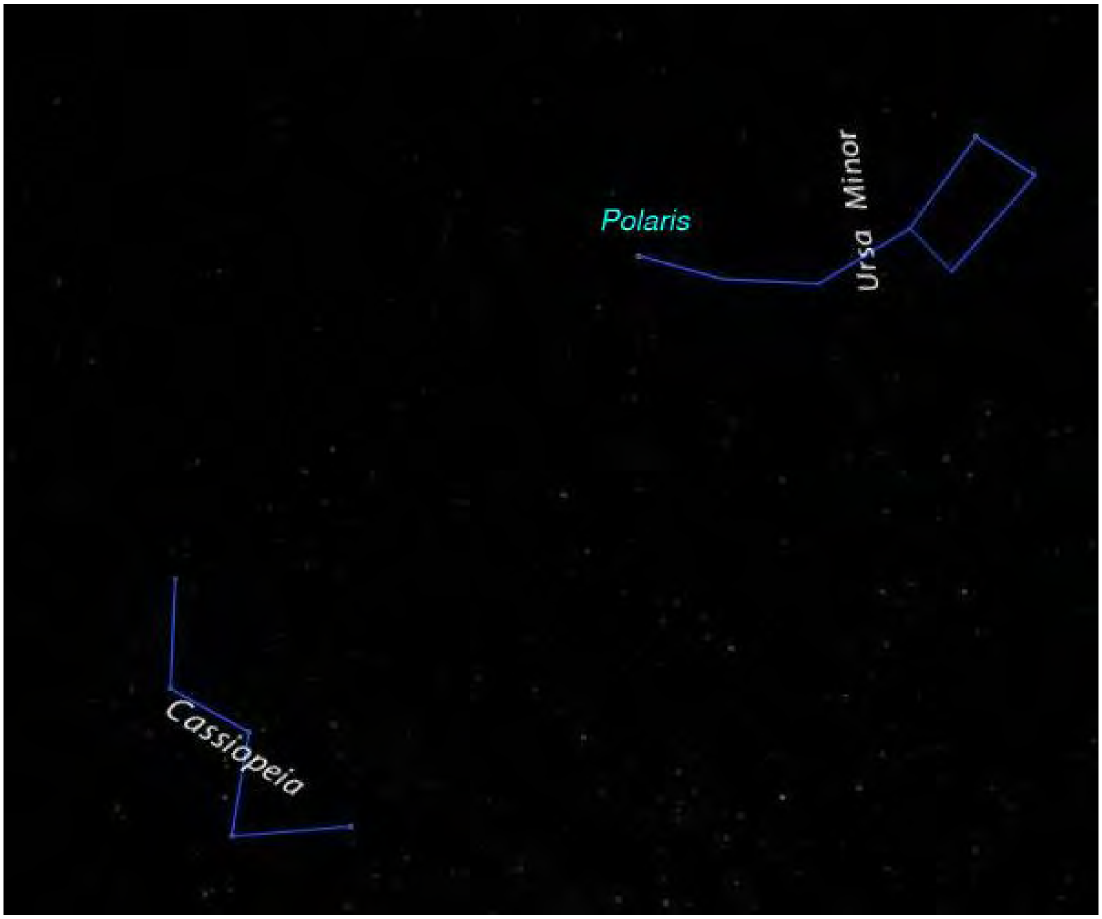


Figure 1.2: Cassiopeia

Roughly the same distance on the other side of Polaris from the Big Dipper is the constellation of Cassiopeia.

or telescopes[2]. In the east, and roughly north of Cassiopeia and Andromeda, is the constellation of **Perseus**, which has a trapezoidal body, and arms and legs sticking out. If we go back to Andromeda, we see that the tip of the letter V is the star Alpheratz which is also part of the constellation of **Pegasus**, the flying horse. The rest of Pegasus is made of a large square, of which **Alpheratz** makes up one corner. The Great Square of Pegasus is especially easy to see because the stars that make up its corners are bright, and only very faint stars are found inside the square.

Already we have enough constellations to tell a story. If you remember your Greek mythology (or perhaps you have seen the movie *Clash of the Titans*), Cassiopeia, the queen of Ethiopia claimed that she was more beautiful than the sea nymphs, the

[2]The Andromeda Galaxy is located roughly 1.5 million **light year**s away, meaning light from it has taken 1.5 million years to travel to our eyes.

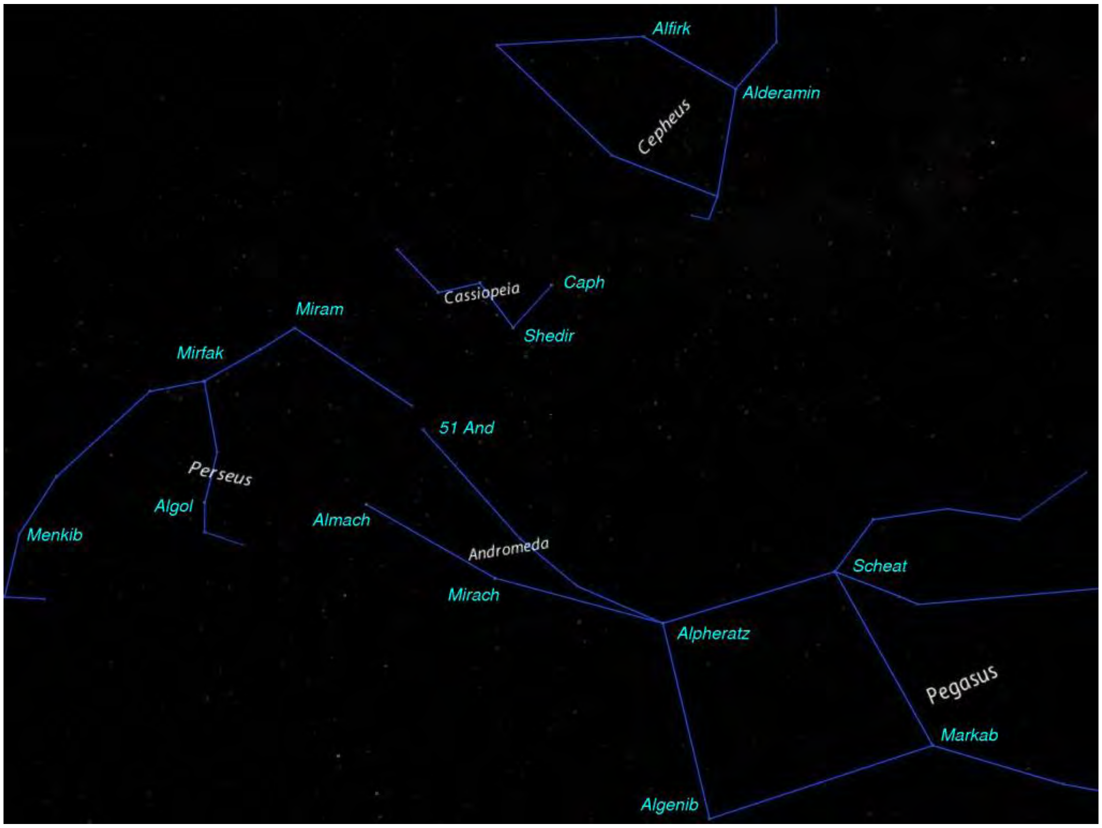


Figure 1.3: Cassiopeia, Cepheus Perseus, Andromeda, and Pegasus

Nereids. This greatly angered the sea nymphs, who complained to their father, the god Poseidon. (In Greek mythology, gods and goddesses were easily insulted, and many stories in Greek myth have the heroes trying to deal with the consequences of their anger). As punishment, Cassiopeia was forced to sacrifice her daughter Andromeda by chaining her to rocks by the coast, where she would be devoured by the sea monster Cetus. However Perseus was traveling through the region, saw Andromeda imprisoned against the sea cliff and fell in love with her. Perseus bargained with her parents that he would receive her hand in marriage if he could defeat the monster. On a previous adventure, he had slain the creature Medusa, who you might recall had slithering snakes instead of hair, and whose gaze was horrific enough to turn to stone anyone who looked upon her. Perseus had cut off Medusa's head, and he turned the decapitated head to Cetus, who immediately solidified into rock. Now Pegasus the flying horse is not in the original story at all, but the filmmakers behind *Clash of the Titans* also managed to work him into the script. In addition to the constellations that we have just covered, in Perseus, you will also find the star **Algol**, which comes from the Arabic for "The Ghoul." This star is often associated with the head of Medusa in that constellation.

1.2.2 The Winter Sky

For people in the northern hemisphere, probably the single most recognizable constellation after the Big Dipper is that of **Orion**, the Hunter. Orion is distinguished by three bright **Belt Stars**, the bright red giant star Betelgeuse at its left shoulder, the bright blue star Rigel at its right knee, **Orion's Sword** that hangs down from the Belt Stars, and the recognizable shape of a person usually viewed with one arm extended outward, and the other arm raised. Orion fell in love with the goddess of the hunt, Artemis. The god Apollo did not want his sister to fall for this mortal, so he sent a giant Scorpion to chase after Orion, who jumped into the ocean to escape its sting. Apollo then tricked Artemis into shooting an arrow into the sea, where Orion was hidden, and thus accidentally killing him. She tried to have the physician Aesclepius save Orion, but he was not able to do so. In the end, she had Orion placed up in the stars. The Scorpion was also transfered into the sky as the constellation of **Scorpio**. Scorpio is on the other side of the sky, so that Orion rises when Scorpio sets and Orion fades from view when Scorpio appears. So even today, the Scorpion continues its chase of the Hunter.

To the southeast of Orion is the brightest star in the sky, Sirius. Sirius is the brightest star in the constellation of Canis Major, the big dog, and hence is also referred to as alpha Canis Major. In sky lore, this constellation is usually referred to as one of Orion's hunting dogs. Sirius was especially important to the Egyptians since its first appearance in the morning sky marked the flooding of the Nile and the beginning of the Egyptian calendar (see § 1.4). Because this occurs in the summer months (today in late June, but for Egypt 3000 years ago, it occurred in early July), Sirius was thought to contribute to heating the Earth, and making the summers hot. Thus our expression "the dog days of summer" actually has an astronomical connection to the Dog Star.

Directly below Orion is Lepus, the Hare, which some accounts claim was placed in the sky for Orion to hunt.

To the west of Orion is Taurus the Bull which is most visible by the stars that make up its long pair of horns. The face or head of the bull is represented by a V-shaped cluster of stars with the bright red **Aldebaran** usually representing the eye of the bull in pictures. The constellation has been identified with a number of different possible bovines, but a story that is very common today is its association with Zeus. The king of the gods fell in love with the Phoenician princess Europa, and he turned into a gentle, white bull to get close to her. Zeus then carried her off, crossing the Mediterranean to get to Crete, where he revealed himself to her. There Europa became the first queen of Crete, and had three sons by Zeus, including Minos, the future king of Crete. This story also provides a mythical answer to why Cretan culture regarded the bull so highly.

The middle of the V contains a cluster of stars, the Hyades, that is easily seen in binoculars. Above the back of the Bull are the **Pleiades**, the Seven Sisters, probably the most famous and most easily seen cluster of stars in the sky. The Hyades and the Pleiades were each a group of seven daughters of the titan Atlas (by different mothers). The Japanese term for the Pleiades is Subaru, and if you look closely at the next Subaru vehicle that you pass by, you will find that the car company uses the stars in its logo.

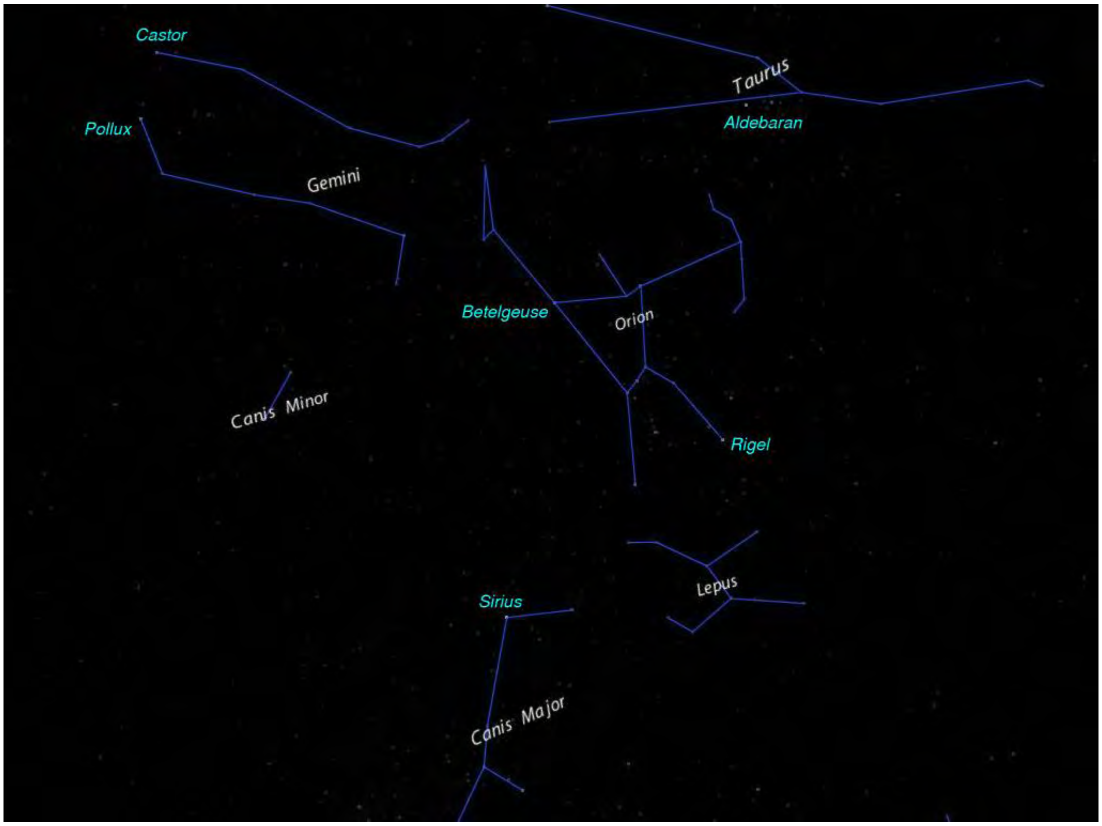


Figure 1.4: The Winter Sky

To the northeast of Orion are **Gemini**, the Twins, anchored by the two bright stars, **Castor** and **Pollux** (with Castor being the most northern twin). Both were the sons of Leda, who was seduced by Zeus after he turned into a swan to get close to her. Castor was the son of Zeus which made him immortal. The two brothers were extremely close and had many adventures together. Their last one was as members of the Argonauts, who with Jason were on the campaign for the Golden Fleece. Pollux was killed during the quest, and Castor became inconsolable from the loss, and petitioned his father so that he could die as well. Zeus was touched by this act of filial love and placed both brothers in the sky.

To the south of Gemini is the star **Procyon**, part of **Canis Minor**, the Lesser of Orion's two hunting dogs.

Further to the east of Gemini is **Cancer**, which consists of faint stars so is very difficult to see from any location with light pollution. The constellation is usually marked by four stars, with one inside the triangle formed by the first three.

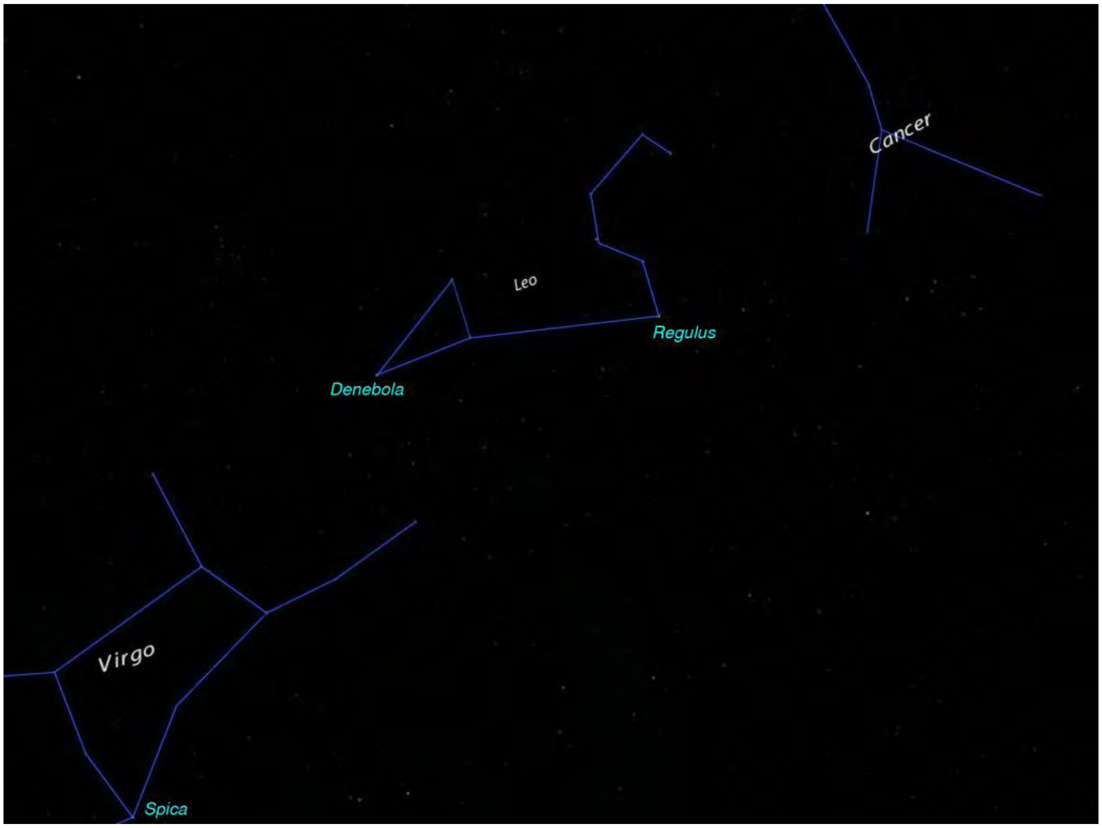


Figure 1.5: The Winter Sky: Leo and Cancer

In late winter, it is easy to find **Leo** the Lion, up in the night sky without having to stay up too late. The shape of a reclining lion, with a maned head is easy to imagine from the bright stars of the constellation. **Regulus** is its brightest star and is located at its front paw; the second brightest is **Denebola**, literally meaning the *Lion's Tail* from the Arabic.

Now form a line between Denebola (at the tail of Leo) and the star **Zozma** at Leo's back near where his rear leg should be. Extend this down to the southeast and it will connect with **Spica**, at the southernmost edge of **Virgo**, the next constellation in the zodiac past Leo. Virgo is a large, sprawling, and dim (except for Spica) constellation, usually depicted as a figure of a woman, oriented sideways in our sky, with her head closest to the Lion.

1.2.3 The Spring Sky

Start at Ursa Major, the Big Dipper, which should be high in the early evening sky. Find the end of the Dipper and follow the arc of the handle until you "arc" over to **Arcturus**, the brightest star in the constellation of **Boötes**. Boötes is easy to identify for modern day skygazers because the constellation is shaped like a giant ice cream cone, with Arcturus located at the tip of the cone. Boötes is usually identified as the Herdsman in the sky, watching over his oxen—not Taurus the Bull, which is halfway across the sky, but the oxen pulling the wagon or wain that is associated with the Big Dipper by many northern European peoples. However Boötes has also been called the "Bear Watcher," by those cultures that saw a bear in that grouping of stars. In fact the name of its brightest star, Arcturus, has been applied to the constellation as a whole, and comes from the ancient Greek words for "Bear Guard."

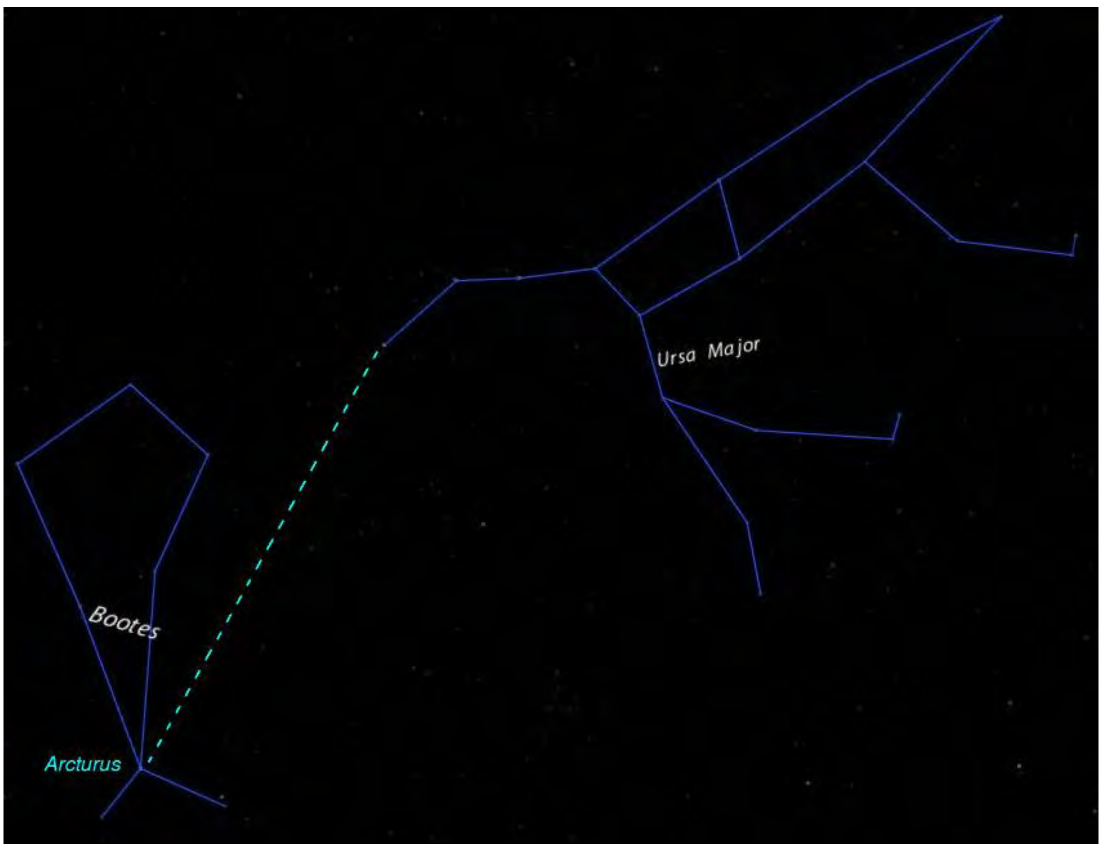


Figure 1.6: Arcing to Arcturus
"Arc" to Arcturus by following the handle of the Big Dipper to the bright star Spica.

In the southern sky, look for the bright red star **Antares**, which is in the eye or head of **Scorpio**, who chases after Orion the Hunter. The Scorpion's tail arches down and to the east away from Antares, and is also easy to spot. However it is located just above the southern horizon from Denver so if there is sufficient light pollution to the south, you might miss it. The claws of the scorpion extend westward and join up with stars that make up **Libra**, the Scales.

Figure 1.7: Scorpio

1.2.4 The Summer Sky

The summer sky is dominated by three bright stars that form the **Summer Triangle**. They are **Vega**, **Deneb**, and **Altair**. Vega is the brightest of the three, and is located in the small constellation of **Lyra**, the Lyre. Deneb can be found in the enormous northern cross that makes up **Cygnus**. Recall that *Denebola* was the Tail of the Lion. Here, Deneb is the tail of the Swan, with its head at the end of a long neck on the opposite side, and the wings extending outward to form the arms of the cross. Altair is the eye of the Eagle, **Aquila**. The body of the bird lies to the south of that star, with the wings extending (appropriately enough) spread-eagled to the sides.

Toward the southern horizon, and to the west of Scorpio, is another easy-to-spot constellation. **Sagittarius**, the Archer, has the appearance of a large teapot that includes a spout (on the west end), a handle (to the east), and a cover (to the north). Sagittarius is important to astronomers because it is in its direction where the center of our Galaxy can be found (something that will be taken up in Lecture 5).

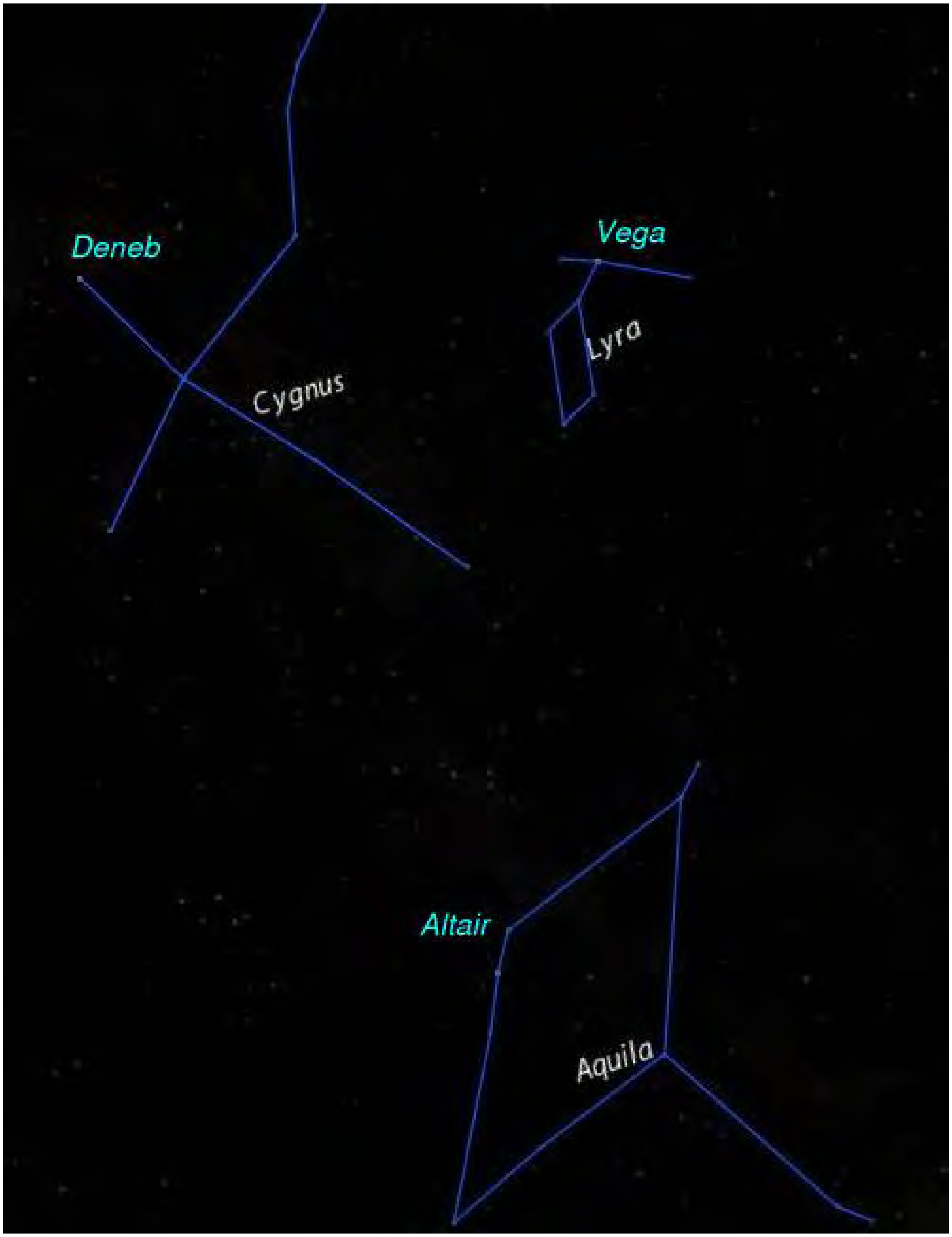


Figure 1.8: The Summer Triangle

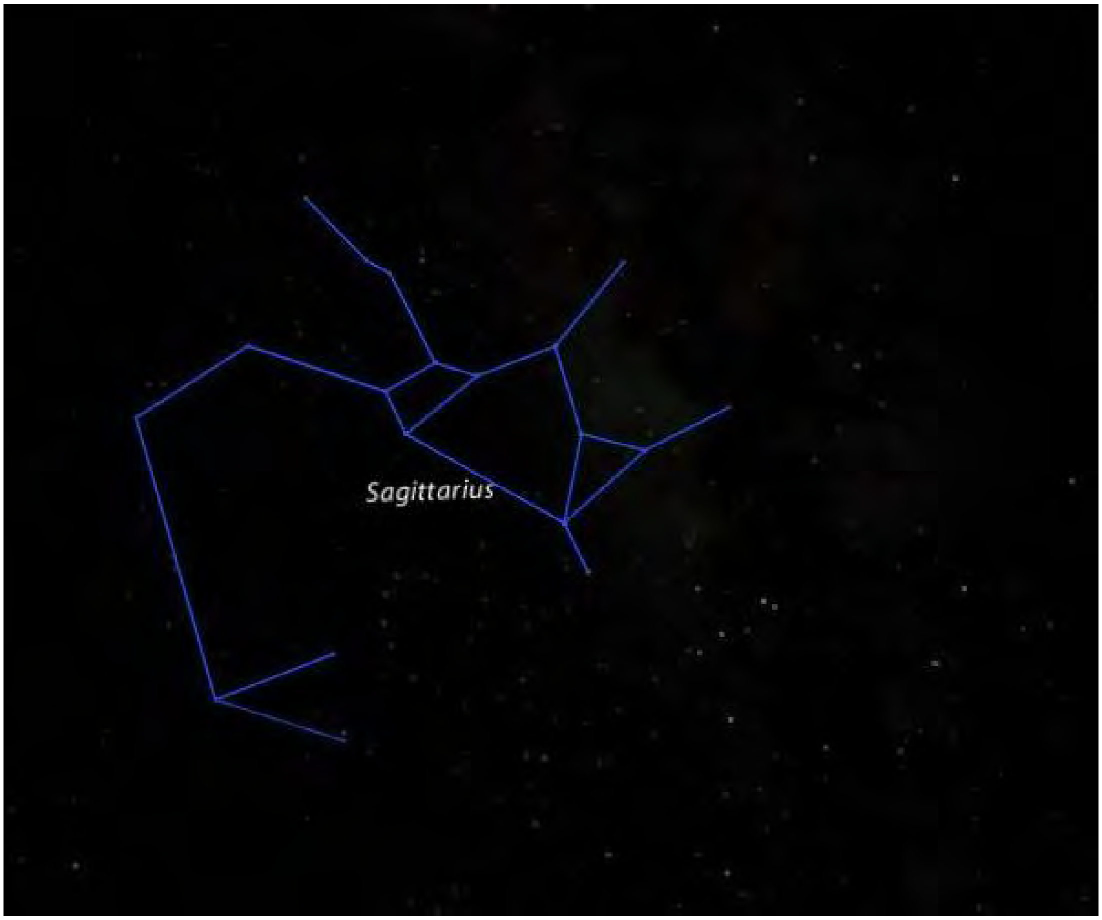


Figure 1.9: Sagittarius

As some of the previous accounts show, constellations can be very useful for learning about the stars. You can assign constellations characters from your culture's myth or other stories that you can easily remember. Although many of our constellations are named after figures in Greek mythology, the Big Dipper, also known as **Ursa Major** or the Great Bear, was identified as such in many cultures throughout Europe, Asia, and North America. To connect us to the Museum's current exhibit, *Quest for Immortality*, this is true even for the ancient Egyptians, which is an curious fact because bears were never found that far south. As a result some astronomers have suggested that the bear identification originated in Eurasia during the Ice Ages and was disseminated by various peoples through the millenia as they migrated and spread out to different continents. And although the Greek myths are prominent in today's accepted identifications of constellations, do not think that other cultures haven't played a role. Our zodiacal constellations originate from the civilizations in the Near East; for instance 5000 years ago, Leo and Taurus made frequent appearances in

Mesopotamian artifacts.

1.3 The Stars from Different Latitudes

We have discussed how the night sky and the constellations appear to us here in Denver. What if we were to go to Minneapolis instead? The stars would so that Polaris would be higher up in the sky, closer to the **zenith**, the direction directly above our heads.

If we were to move further north say to Anchorage, Alaska (at a latitude of 61°), we see Polaris shift even more. At this location, the motions of the stars trace more complete circles in the sky. Only stars furthest to the south have incomplete arcs.

Similarly if we physically travel south toward the equator, the North Star will move lower in the sky toward the northern horizon. In this planetarium, Polaris will move beyond the edge of the back of the visible dome. The majority of the stars now clearly rise from the east and set in the west.

How can we explain what is going on in the sky given our modern day knowledge of the Earth? Luckily we all know today that the Earth is not flat, but is spherical, giving us a leg up on the ancients who did not have this knowledge. We also know that the Earth rotates on an axis; that the intersection of this axis on the surface is geographically labeled as the **north pole** and **south pole**, and the great circle line equidistant between the two is the **equator**. The Earth can be further divided into lines that mark increasing **latitude**s north and south of the equator, as well as **longitude** lines east and west of some arbitrary point[3]. The reason why the stars appear to rise and set in our sky is not because they are fixed to the sky, and that is rotating above us. Instead, the Earth is rotating underneath the fixed stars.

We can visualize this more clearly if we take the north and south poles of the Earth, as well as its equator, and extend them outward until they connect with the celestial sphere of the sky. *These intersection points mark the celestial north and south poles, as well as the celestial equator.*

If you have ever sat on a train, and it starts to slowly move in the forward direction, it will seem to you that the world is moving in the opposite, backward direction. This will be the case especially if the motion is so slow that you don't feel yourself moving. Similarly, as the Earth rotates on its axis during the course of a day, to us sitting stationary on the ground, the sky will appear to turn in the opposite direction.

A spherical Earth spinning on its axis also explains our different views of the heavens as we move in latitude. When we are further south, closer to the Earth's equator, the celestial equator will be high in the sky, and the celestial North Pole (and hence Polaris) will be further toward our northern horizon. The stars will tend

[3]This arbitrary **meridian** line of 0° longitude is defined in modern times to run through Greenwich, England. By our modern convention of how the Earth is subdivided, the geographic coordinates of Denver is roughly 105° West, and 40° North.

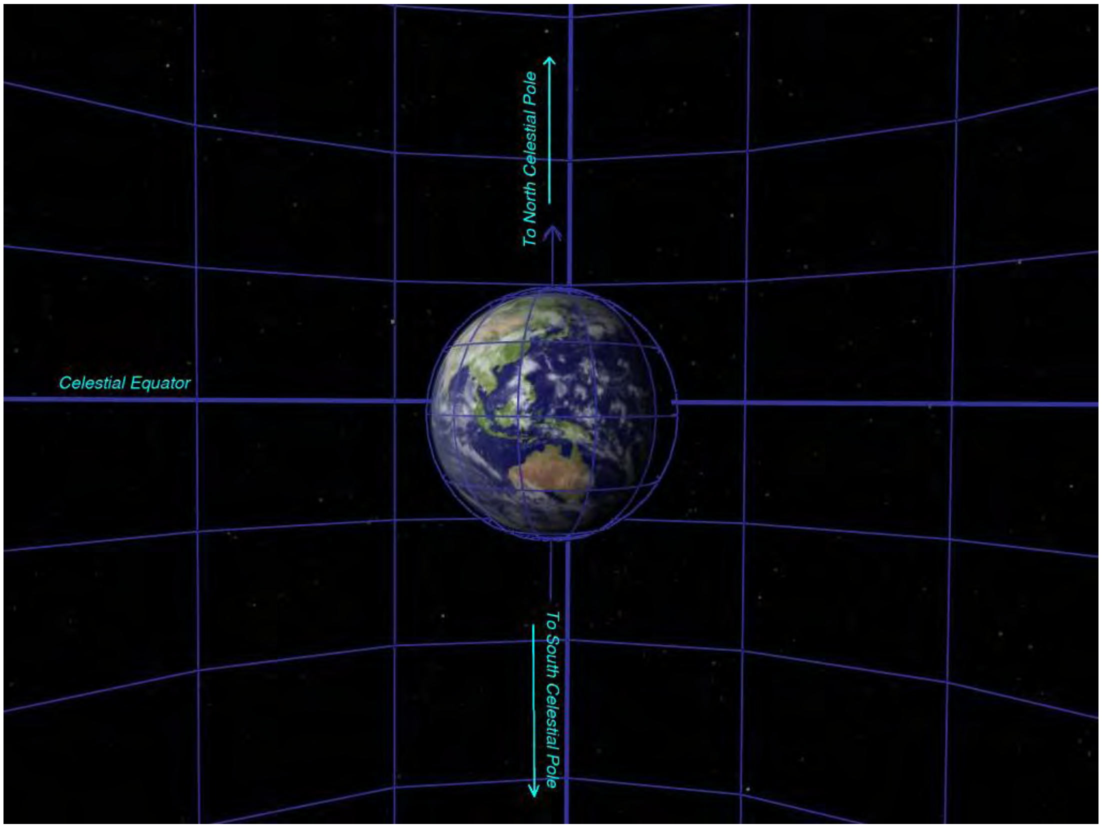


Figure 1.10: The Celestial Equator and Celestial Poles
Extend the Earth's equator and poles out into infinity, and where they intersect the celestial sphere are the celestial equator *and* celestial poles*.*

to more clearly rise, arc, and set.

If we transport ourselves on the Earth to the northern latitudes, Polaris will move higher up in our sky. The celestial equator moves down toward the horizon; and the stars now move in large horizontal arcs between rise and set. If instead of watching the stars, we watch the Sun, we also understand why during summers at high latitudes near the Arctic (or Antarctic) circle or above, the Sun can stay up in the sky for so long. Instead of rising and setting overhead us, it skirts along the horizon, dragging out the length of the day.

1.4 Star Motions Through the Year

We have seen what constellations appear in particular seasons. Over the course of a year, some of the constellations disappear and new ones appear. This is because from day to day, any given star rises and sets slightly earlier each night.

Let's say you observe a star in the eastern horizon that is just barely visible in the glow of the rising Sun. On the day before, the star will have risen just a few minutes later, and hence the sky will be too bright for the star to be seen. On the day after, the star will rise a few minutes earlier and will be in the sky a few minutes longer before the Sun rises.

Figure 1.11: September Night Sky
A northern hemisphere person looking out into the night side in September (standing on the back side of the Earth shown here) will see constellations like Pegasus and Andromeda high in the sky.

If you compare the daily motions of the Sun with the daily motions of the stars, they are *not exactly the same!* Whereas the Sun takes 24 hours (a *solar day*) to make one complete circuit around our sky—say from noon to the next noon—the stars actually take slightly less time, about 23 hours 56 minutes (or a *sidereal day*) instead.

The day on which a star is first visible in the eastern sky right before dawn is called the day of **heliacal rising**. Similarly the day on which a star is first visible in the western sky right after the setting Sun is known as the as the day of **heliacal setting**[4]. Stars will appear in the sky at such heliacal rising and settings at exactly the same date every year. This is an important date-keeping tool for those without access to clocks or calendars. Ancient peoples used such appearances to tell them what date it was, within an accuracy of a day or two. Again to bring this back to

[4]The word *heliacal* means "near the Sun."

Figure 1.12: September Day Sky
At the same time of the year, the Sun will appear to be in the constellation of Virgo.

Quest for Immortality, the heliacal rising of Sirius, the brightest star in the sky, was used to begin the Egyptian calendar, since this marked the annual flooding of the Nile. Still other peoples used other stars to mark the beginnings in their calendar, or the start of the growing season; this has included the Belt Stars of Orion, and the Pleiades which have been important for Australian aborigines as well as for some Native American tribes.

Because of this difference between the daily motion of the stars and the Sun, if you watch the night sky over time, say over many weeks and months, you will see different constellations come into view. How do we explain what is going on given our modern day knowledge of astronomy? In addition to the Earth turning on its axis, we also know that it revolves in an orbit around the Sun, and the time for it to make one complete circuit is one year, or $365\,{}^{1}/_{4}$ days.

During the month of September, the Earth is in a position in its orbit so that at roughly midnight, a person will see Pegasus and Andromeda high up in the sky; if we were to look in the other direction (and if we could see the stars amid the glare of the Sun), we would see the Sun in the constellation of Virgo.

Over time, the Earth moves in its orbit. So if we were to wait a few months, and check the sky at Christmas, the Sun would be in the constellation of Sagittarius. At night, we would see a completely different set of constellations up in the midnight sky, including Gemini, Cancer, and Orion.

The small four minute difference between a star finding itself back in the same position in the sky versus the Sun getting back in its original position is directly related to the Earth moving in its orbit. Each day, the Earth moves a little bit further in its orbit. It is enough that the Earth has to rotate an extra four minutes before a person on the Earth can see the Sun in the same position in the sky.

If we further move through the year, we will see the Sun track through different constellations over the course of the year. In late September through October, it would be in the constellation of Virgo, then move into Libra in November, followed by Scorpio, cross the tail of Ophiuchus, and then Sagittarius around Christmas and through the middle of January, then into Capricorn and Aquarius, then Pisces, Aries, Taurus, Gemini, Cancer, Leo, and then back to Virgo again one year later.

The line that the Sun follows in the sky through the course of the year is the **ecliptic**. It is defined by the Earth's orbit. The path of the ecliptic also defines a region in the sky known as the **zodiac**. Hence nearly all of the constellations that the Sun moves through are known as the zodiacal constellations.

1.5 The Seasons

The ecliptic is tilted by with respect to the Earth's equator and poles. This tilt is 23.5°. (When you look at a globe, you will now know why it is tilted and what this tilt is with respect to!) The tilt of the Earth with respect to its orbit, the ecliptic, also explains another phenomenon: the **seasons**.

On a date in late June, the Earth is oriented in such a way that the northern hemisphere is tilted toward the Sun. In fact, the line dividing day and night, the **terminator** cuts across the globe on the opposite side of the North Pole. People living above the Arctic Circle on this date would never see the Sun set. People living in the north not only get more sunlight during the day, but the light falling upon the surface is more direct, hence warming the northern hemisphere more. By contrast, the southern hemisphere gets much less light. As a result, the northern hemisphere experiences summer while the southern hemisphere experiences winter.

Three months later, the tilt of the Earth hasn't changed, but it has moved in its orbit so the alignment of the Earth with the Sun brings equal illumination over both northern and southern hemispheres.

Now if we wait another three months (hence, it is six months after we started), the Earth will be on the other side of the Sun. The southern hemisphere is getting not only longer but more direct solar illumination. Now summer will have started for the southern hemisphere, while the northern hemisphere moves into winter.

Three more months later and both hemispheres are now equally lit again. We are at the start of spring, with the north warming and the south cooling. If we move forward another three months, we are back to where we started and we are again at the start of the northern summer.

(a) Sun in Libra

(b) Sun in Scorpio

(c) Sun in Sagittarius

(d) Sun in Capricorn

(e) Sun in Aquarius

(f) Sun in Pisces

Figure 1.13: The Sun at Different Times of the Year
As the Earth moves in its orbit, the Sun will appear to be located in different constellations during different times of the year. These particular constellations are part of the zodiac.

1.6 The Phases of the Moon

One object we have seen tonight but have not explicitly discussed is the Moon.

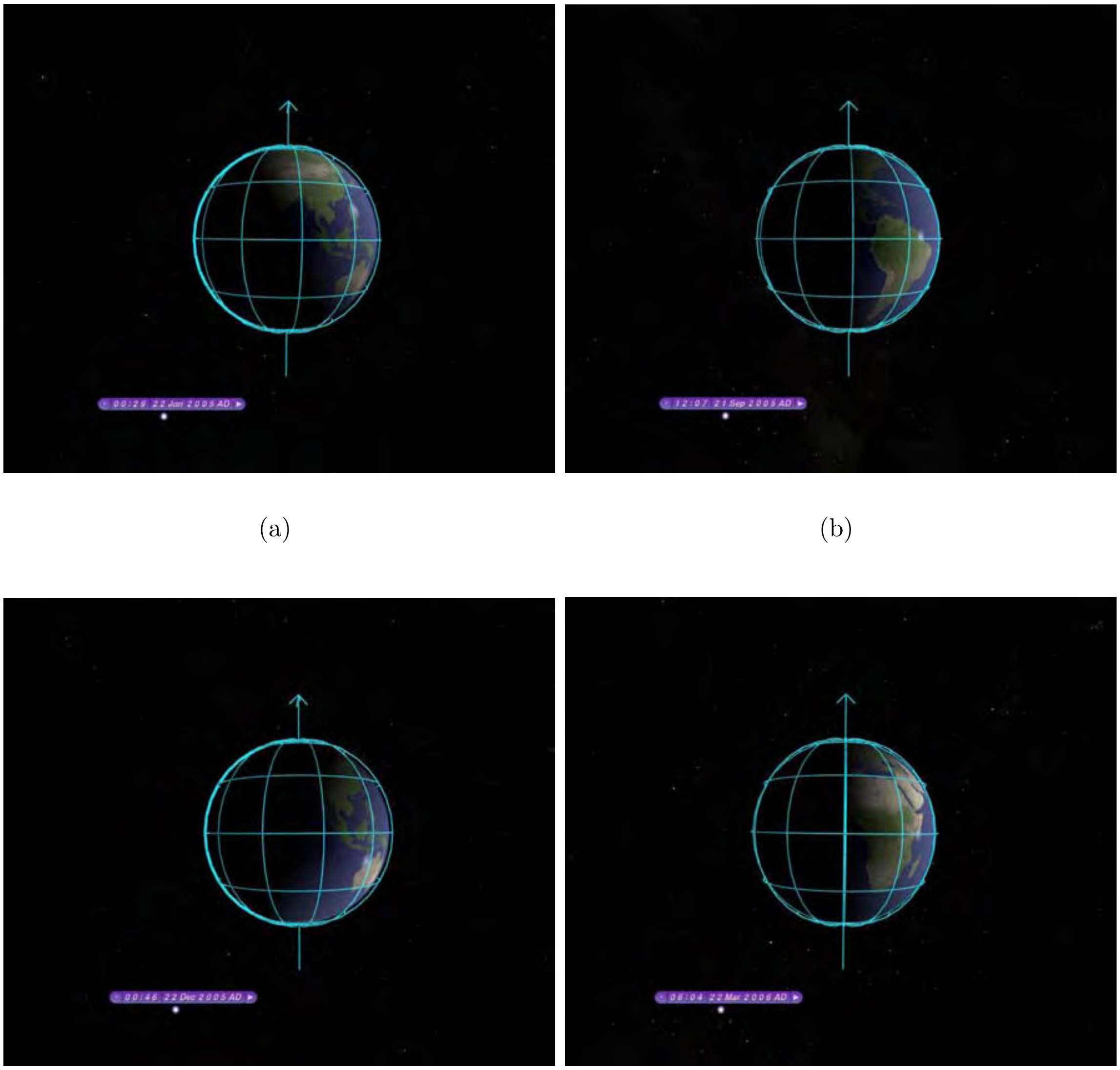
Let us lock to the Moon, so that the Earth is behind us. Tonight the we see a nearly **Full Moon**, but if we move forward a week, we will see the Moon wane to **3rd Quarter**. Notice the reason why it appears that only one side is lit is because of its orientation with the Sun.

Now we move forward another week and we see the Moon has less of its surface illuminated by the Sun. It is getting to the point of what is called **New Moon**.

Notice that for the New Moon, the Moon is between us and the Sun. And depending on the orientation of the Moon's orbit, the Moon can pass quite far from the Sun, when it is new, or in next month's case, pass quite close.

Finally if we lock back onto the Moon, and move another week forward in time, we find ourselves with a **1st Quarter** Moon. And yet another week later, the Moon continues to grow or *wax* until we are back at Full Moon. Here the Moon is opposite the Sun, on the other side of the Earth, and we see its fully illuminated face.

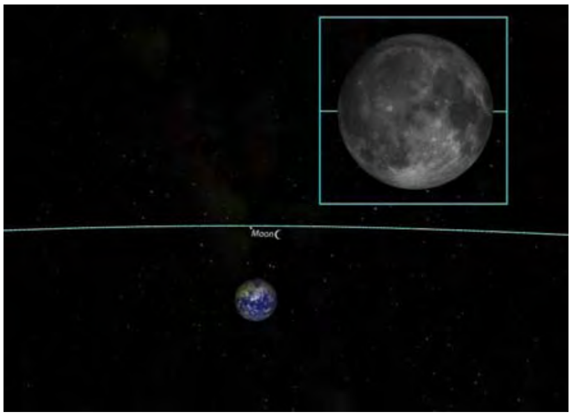
While the Moon orbits the Earth over the course of a month, its phases smoothly change. This is an effect entirely due to the relative orientation of the Moon with respect to the Sun. If we watch the Moon over its next orbit, we can tell what its phase will be depending on its orbital position: first starting with Full, then 3rd Quarter, then New, then 1st Quarter, and then back to Full again.



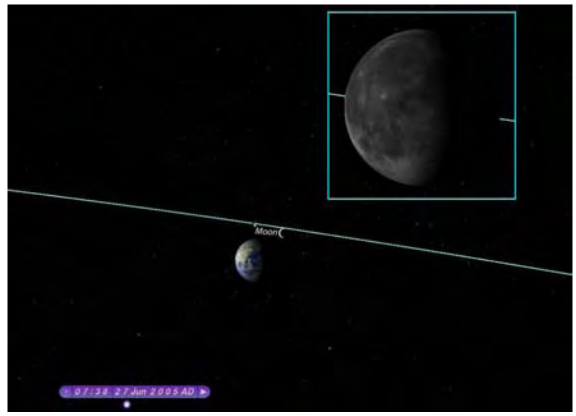
(a) (b) (c) (d)

Figure 1.14: The Seasons on Earth

Because the Earth is tilted with respect to its orbit, the amount of sunlight received by the northern and southern hemispheres vary throughout the year. In late June, the terminator *line dividing night and day leaves more of the northern hemisphere exposed to sunlight. During late September, both hemispheres receive equal amounts of Sun. Near the end of the year, the southern hemisphere gets the bulk of the light and heat. Finally in late March, the northern and southern hemispheres again reach parity.*

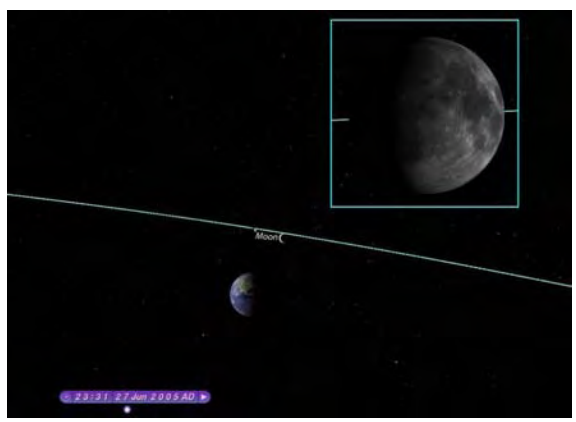


(a) Full Moon



(b) 3rd Quarter

(c) New Moon



(d) 1st Quarter

Figure 1.15: Phases of the Moon

Depending on the location of the Moon with respect to the Sun, the Moon's appearance will appear to shift through a series of phases. *A Full Moon is opposite the Sun from the Earth; a New Moon is in the same direction as the Sun. The 1st and 3rd quarter Moons are intermediate between the two.*

Lecture 2

The Terrestrial Planets

2.1 The Ancient View of the Planets

In addition to the stars in the nighttime, which were fixed and unchanging, ancient peoples also noticed six points of light, which not only tended to be brighter than the other stars, but moved between the stars over time. The ancient Greeks called them the "wanderers" and it is their word for that term that has been passed on down to us today as the **planets**.

Figure 2.1: Aristotle and Ptolemy

Five planets were visible to the ancients. All of them moved in ways similar to the motion of the stars in the celestial sphere. Hence they were seen by most ancient thinkers as circling the immobile Earth that was located at the center of the universe. The Greek philosopher **Aristotle**, who lived in the 4th century BC,

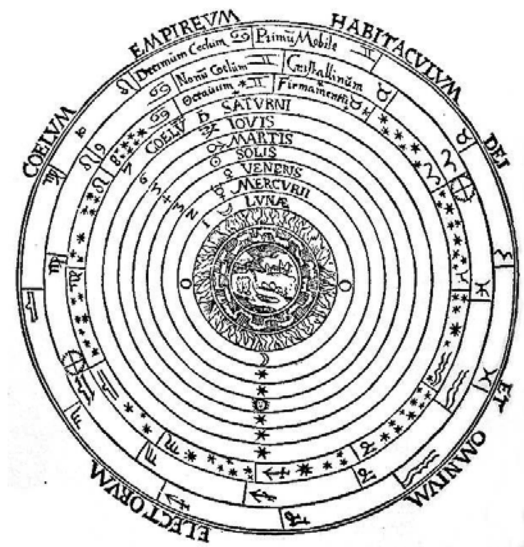


Figure 2.2: Ptolemy's Geocentric Universe
A reproduction from a page of the Almagest*, showing a universe centered about the Earth. In successively more distant spheres are the Moon, Mercury, Venus, the Sun, Mars, Jupiter, Saturn, and the stars in the firmament. The* Primu Mobile *is found in the last and tenth sphere; this "Prime Mover" or central Deity sets in motion everything else in the Universe.*

was one of the chief proponents of an Earth-centered universe. The Sun, Moon, and planets were located on transparent spherical shells that rotated uniformly around the Earth, creating the apparent motions of the celestial objects over time. Aristotle also espoused the view that the Earth, corrupt and imperfect, was composed of four elements (earth, water, air, and fire); the heavens on the other hand were perfect and hence was made of an entirely different element, the *aether*. Although the Sun, stars, and planets were perfect heavenly objects closest to the gods, the Moon being so close to the Earth was tainted by its imperfection, which explained its blotchy appearance.

The Greek astronomer **Claudius Ptolemy** (who lived and worked circa 140 A.D.) wrote a 13-volume work on astronomy, which was passed on to the Arabs after the destruction of the Library of Alexandria. His magnum opus is known to us today by its Arabic name, the *Almagest*, and was accepted as a reference for more than a thousand years. Ptolemy's cosmology had the Earth at the center, with the Moon, Mercury, Venus, the Sun, Mars, Jupiter, and Saturn located on successive transparent shells. Although the Greeks believed the perfect heavens should have motions that followed the perfect geometric shape, the circle, they were not able to model the motions of the planets with simple circular orbits. Instead Ptolemy envisioned large circular *deferents* that surrounded the Sun. The planets themselves moved on smaller circles, the *epicycles*, whose centers moved along the deferent.

With improved observations over time, it became clear that even deferents with

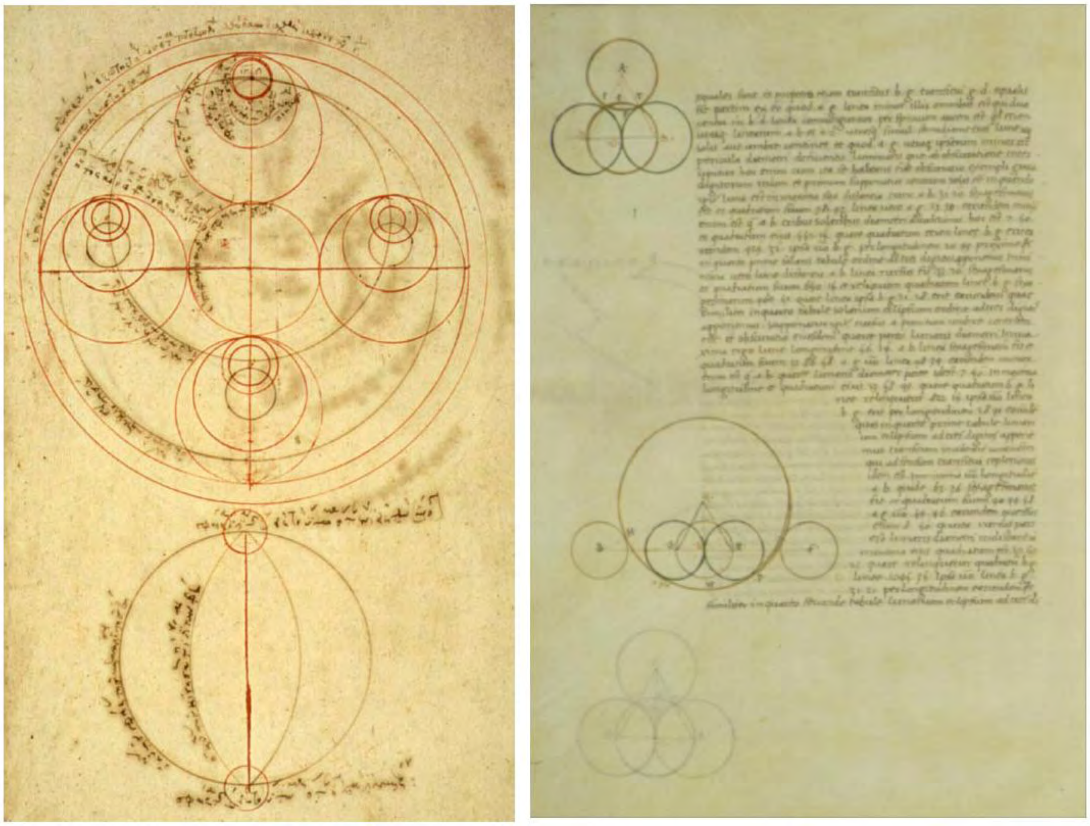


Figure 2.3: Epicycles and Deferents
Two more pages from the Almagest *showing the complexity of a geocentric universe model with only circular motions. Celestial objects move in circular epicycles which themselves move in larger circular deferents.*

a single epicycle was not enough to match the motions of the planets in the sky. Astronomers up through the Middle Ages added increasing numbers of epicycles, "wheels within wheels," to try and match the irregular motions of the planets and our Moon.

2.2 The Copernican Revolution

The Polish astronomer **Nicolaus Copernicus** (born 1473) was worried by the inaccuracies of the Ptolemaic model. Observations showed that Mercury and Venus never wandered very far from the Sun: Mercury was never further than 28° away, while Venus was always within 47°. The set of nested epicycles that were needed to keep these two planets close to the Sun were getting increasingly baroque. Copernicus made the realization that planetary motions could be much simplified if the Sun was moved to the center of the of the Universe as opposed to the Earth. Although Copernicus still needed epicycles to make his planetary motions match the real motions in the sky, the publication of his work, *De Revolutionibus* (*On Revolutions*, 1543), which explained his idea and the evidence for it, shook the philosophical world. Virtually all serious thinkers in medieval Europe placed the Earth at the center of the universe, and Copernicus was attacked by both Catholic Church officials and by leaders of the Protestant movement.

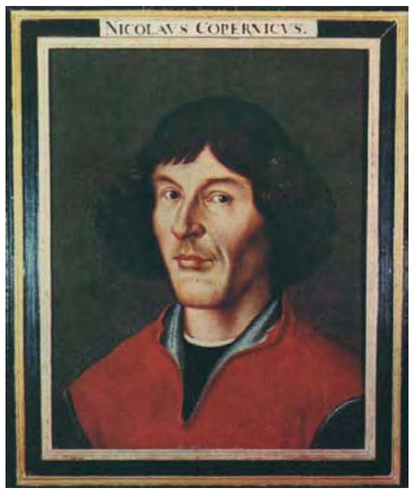


Figure 2.4: Nicolaus Copernicus

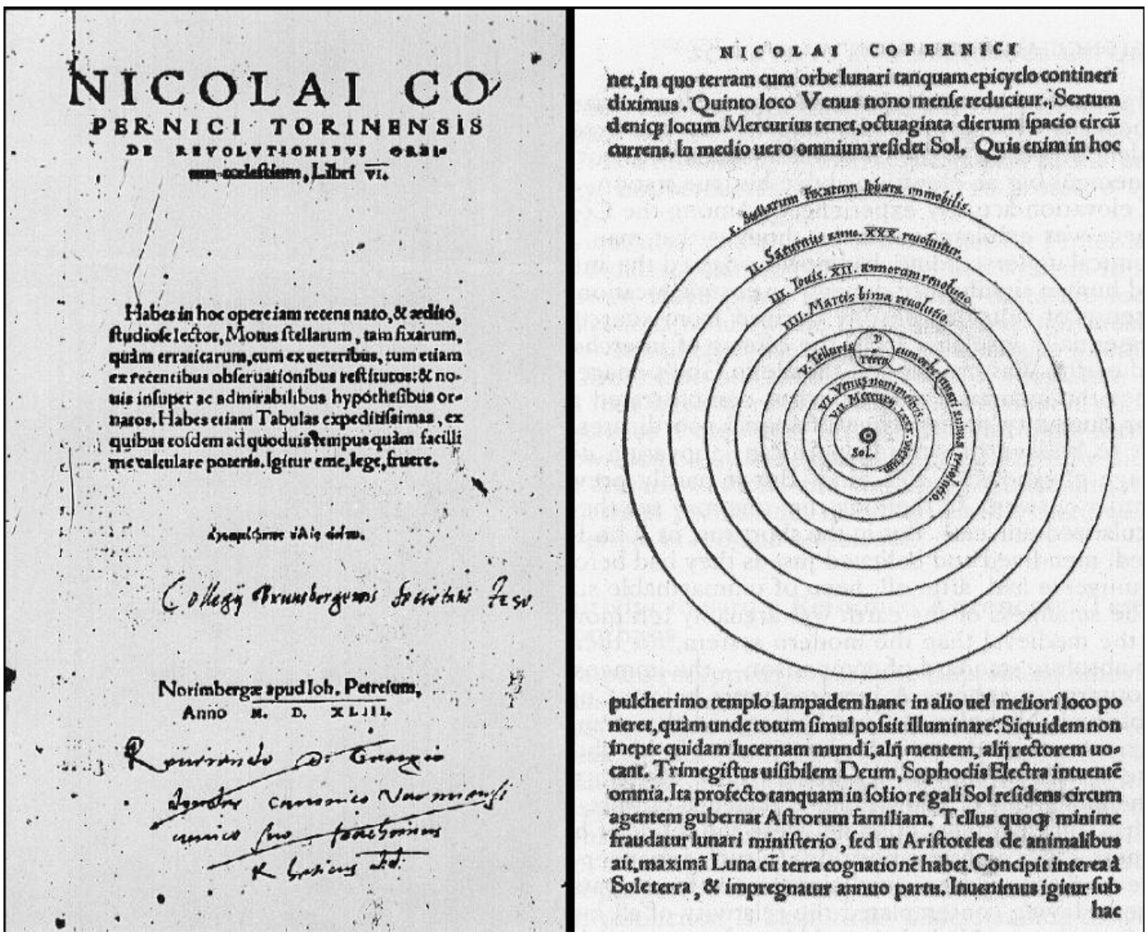
Copernicus died the year that his book was published, and so missed the turmoil that was to ensue from the intellectual revolution that would eventually be named after him. In addition to removing the Earth from the center of the universe, Copernicus postulated further that the stars were immeasurably far away. Other thinkers expanded on this idea. In 1584, the Italian naturalist, Giordano Bruno, wrote that the stars were actually distant Suns, with intelligent alien races that inhabited the planets that circled them. Bruno was burned at the stake by the Inquisition in 1600 for heresy.

In 1609 and 1619, **Johannes Kepler**, published in two books a theory in which he could explain the celestial motions with a heliocentric theory, but he threw out the idea of epicycles entirely. Instead of moving in perfect circular motions, all the planets (which now included the Earth) orbited the Sun in **ellipses**. He further formulated his three Laws of Planetary Motions:

1. A planet moves around an ellipse with the Sun at one focus.

2. The line between the sun and the planet sweeps over equal areas in equal times.

3. The cube of the semimajor axis of each planet's orbit is equal to the square of its period of revolution.



NICOLAI CO
PERNICI TORINENSIS
DE REVOLVTIONIBVS ORBI-
um coelestium, Libri VI.

Habes in hoc opere iam recens nato, & aedito, studiose lector, Motus stellarum, tam fixarum, quàm erraticarum, cum ex ueteribus, tum etiam ex recentibus obseruationibus restitutos: & nouis insuper ac admirabilibus hypothesibus ornatos. Habes etiam Tabulas expeditissimas, ex quibus eosdem ad quoduis tempus quàm facillime calculare poteris. Igitur eme, lege, fruere.

Norimbergae apud Ioh. Petreium,
Anno M. D. XLIII.

NICOLAI COPERNICI

net, in quo terram cum orbe lunari tanquam epicyclo contineri diximus. Quinto loco Venus nono mense reducitur. Sextum denique locum Mercurius tenet, octuaginta dierum spacio circũ currens. In medio uero omnium residet Sol. Quis enim in hoc

pulcherrimo templo lampadem hanc in alio uel meliori loco poneret, quàm unde totum simul possit illuminare? Siquidem non inepte quidam lucernam mundi, alij mentem, alij rectorem uocant. Trimegistus uisibilem Deum, Sophoclis Electra intuentẽ omnia. Ita profecto tanquam in solio regali Sol residens circum agentem gubernat Astrorum familiam. Tellus quoque minime fraudatur lunari ministerio, sed ut Aristoteles de animalibus ait, maximã Luna cũ terra cognationẽ habet. Concipit interea à Sole terra, & impregnatur annuo partu. Inuenimus igitur sub hac

Figure 2.5: *De Revolutionibus*

The title page (left) *and an interior page* (right) *from Copernicus' most important work. The Earth has been removed from the center of the Universe and is now merely one of many planets that are circling the Sun. The diagram on the right does not reveal the true complexity of Copernicus' model, since he was still relying on circular motions, and hence was forced to use epicycles and deferents.*

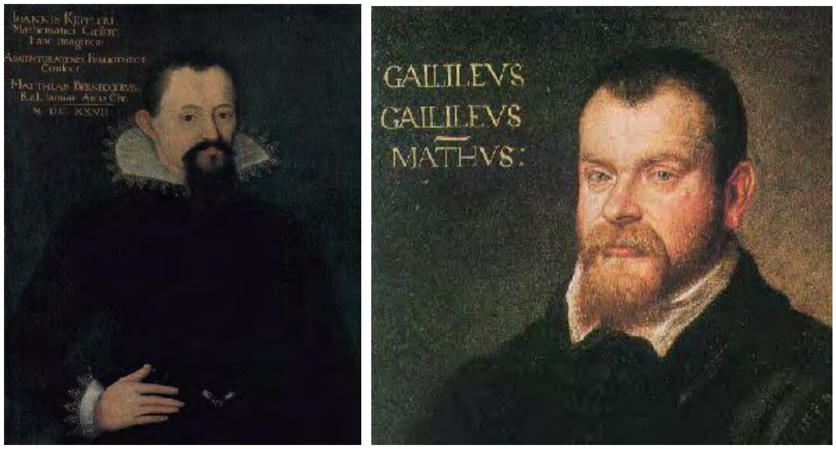


Figure 2.6: Johannes Kepler and Galileo Galilei

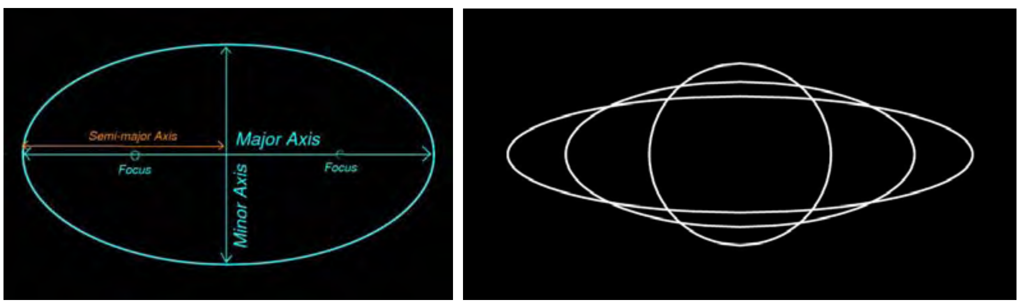


Figure 2.7: Ellipses
The different components of an ellipse (left). *The shape of the ellipse is defined by its major and minor axes. In Keplerian orbits, the Sun is located at one of the two foci. Three ellipses of different eccentricities* (right)*; the circle is actually an ellipse with with zero eccentricity; as the eccentricity is increased, an ellipse becomes stretched out and more cigar-shaped.*

An ellipse is a round, geometric, closed shape, defined by two points called **foci** (plural of **focus**), with a long **major axis** through them, and a short **minor axis** perpendicular to the major axis. A point on the ellipse is defined such that if you were to add the distance from that point to one focus with the distance from the point to the other focus, the sum total would always be the same. Hence one way to draw an ellipse is to place two tacks (or pins) on a sheet of paper to represent the foci; tie a piece of string between the tacks, and trace a closed figure using a pencil placed within the loop of string. An ellipse can have its shape modified by the distance separating the two foci and the relative length of the string compared to this focus distance. Two foci far apart compared to the length of the string will give a flattened, cigar-shape; such an ellipse has **high eccentricity**. On the other extreme, two foci very close together will give an ellipse that looks almost circular, or one that has **low eccentricity**. In fact, a circle is a special case of an ellipse where the two foci have merged together.

The planets in our Solar System have orbits with low eccentricity. Thus the ancients' attempts to model their movements using circles was not far off. However the further into the future that astronomers in the Middle Ages tried to extrapolate planetary positions, the more the actual positions deviated from their predicted positions. The small difference between a circular orbit and a slightly eccentric elliptical orbit added up over time to give large errors to theories with deferents and epicycles. The use of a single ellipse for a single orbit in Kepler's First Law's was a vast improvement over previous theories just based on the simplification it made to the problem.

A direct consequence of Kepler's Second Law is that planets in motion around the Sun will move faster in their orbits the closer they are to the Sun. For a highly eccentric elliptical orbit, the Sun will be at one focus, located near one end of the orbit. Objects with such orbits (such as comets) will whip around the Sun in a fraction of the time it takes for it to move through the opposite side of the ellipse. For orbits that are circular, the motion of the planet will be constant.

Finally the Third Law says that with increasing sizes of orbits, the longer it will take a planet to make one complete circuit around the Sun. There is a precise mathematical relationship between the **semimajor axis** (which is almost the same as a planet's distance to the Sun for most of the planets) and the time or **period** of an orbit.

Let us work with the distances to the planets (since they are almost the same as the semimajor axes for low eccentricity orbits). The distance from the Earth to the Sun is 150 million kilometers, and is defined as 1 **astronomical unit** or 1 AU. We can make a table denoting the distances to the first six planets from the Sun in AUs as well as the cube of the distances (= distance^3 = distance $\times$ distance $\times$ distance):

Table 2.1: Planet Distances

Planet	Distance (AUs)	Distance^3 (AU^3)
Mercury	0.39	0.059
Venus	0.72	0.373
Earth	1.0	1.000
Mars	1.5	3.375
Jupiter	5.2	140.6
Saturn	9.5	857.4

We can also construct a table listing the orbital periods in Earth years, and the square of those periods (= period^2 = period $\times$ period):

Table 2.2: Planet Periods

Planet	Period (years)	Period^3 (year^3)
Mercury	0.24	0.058
Venus	0.62	0.384
Earth	1.00	1.000
Mars	1.9	3.61
Jupiter	11.9	141.6
Saturn	29.4	864.4

With the limited accuracy numbers that we used to create these two tables, we see that the last two columns are nearly identical. Thus the periods and sizes of the planetary orbits follow the equation given by Kepler's Third Law!

One person who bears final mention is the Italian scientist **Galileo Galilei**, who started studying the heavens in the early 17th century, but putting to use a new tool, the telescope. Among his landmark observations were the discovery of phases for Venus. For this to occur, the Ptolemaic model of Venus' orbit must be wrong, since Venus would have to spend part of its time on the far side of the Sun from the Earth, for this to happen. When he turned his telescope to Jupiter, Galileo also discovered its four large moons (which are now called the *Galilean moons* after him). This was evidence that not everything in the heavens revolved around the Earth, which was another attack on the tenets of Ptolemaic cosmology. Finally Galileo's discovery of mountains and other geological features on the Moon, and his observations of spots on the Sun, proved that the heavens were not the pure, objects of perfection that the ancients had imagined. Instead they were individual worlds, that were probably very different from the Earth, but were nevertheless probably composed of the same sort of materials that could be found on Earth.

2.3 Origin of the Solar System

Before covering the specific details concerning the terrestrial planets, it is useful to understand exactly how the planets got to be the way they are now. To do this, we must an understanding of the current theories of the formation of the Solar System.

Based on an enormous amount of observational and theoretical evidence, from planets in our Solar System, the Sun, and meteoritic evidence, as well as astrophysical research on star formation, a slow consensus has been building on how our planetary system arrived at its present form. More details on the star formation process will be given in Lecture 4. The basic story is the following. The Sun and its satellite planetary system originated in a massive cloud of gas and dust, the initial **solar nebula**. Under the effects of gravity, this nebula started contracting more than 4.5 billion years ago. Each arbitrarily clump of matter in the gas cloud has some small amount of **angular momentum** (a term in physics which involves the measure of the spin or rotational motion of that matter). As the cloud proceeds in its collapse, and as clumps of matter collide with each other, angular momentum in one direction in one parcel of gas can be canceled out by an opposite angular momentum of another parcel of gas that is spinning in the reverse direction. Over time, the collapsing cloud will have a **net angular momentum**, whatever is leftover after these multiple cancellations.[1]

Angular momentum is a basic physical property that is conserved. It can never be destroyed or created, only transfered from one physical system to another. As a consequence of this conservation, as matter in the solar nebula collapses inward, it begins to spin around the center faster. The net angular momentum also manifests itself in the shape of the evolving solar nebular: it flattens into a disk with the central proto-Sun at the center.

Gas and dust will continue to spiral into and accrete onto the protostar at the center. As the center heats up from nuclear reactions (some of which begin very early on, perhaps 100,000 years or less after the onset of collapse), and from the heat of collapse, the solar nebula has the highest temperatures right next to the proto-Sun (up to several thousands of degrees Celsius) and dropping off as you get further from the center.

Not all of the gas and dust eventually winds up in the Sun. A small fraction manages to gather together to form the planets. However different elements and compounds have different **condensation temperatures**, the point at which they solidify. With adequate pressure, water, for instance, condenses from a gaseous phase to a liquid phase when the temperature falls below 100° C (212° F). It further changes state and condenses again to ice at 0° C (32° F). Other compounds, like methane, carbon dioxide, and ammonia, are similar to water in having relatively low condensation points. These are **volatiles**, which tend to be formed from small numbers of lighter elements, such as hydrogen, carbon, nitrogen, and oxygen. Other materials are **refractory**: they condense at much higher temperatures and the condensed matter has properties with which we would associate with the solid earth or ground. They can contain heavier elements, such as silicon, phosphorus, sulfur, and calcium; metals like cobalt, nickel, and iron; and radioactive species like uranium and thorium.

Because of the temperature difference between the part of the solar nebula close to the Sun and that part of it far away from the Sun, a **condensation sequence** appeared in the disk. Refractory materials such as silicates and metals condensed out of the nebula gas close in, while ices and gases condensed out much further out. As a result, all of the inner planets close to the Sun tend to have a larger fraction of rocky

[1] Why does the cloud not end up with zero angular momentum? It is very difficult for any macroscopic ensemble of particles in the universe to have exactly zero angular momentum. At least not for very long since by its interactions with the rest of the universe will it gain some net angular momentum. One might as well try to find a collection of particles that is absolutely at rest with itself.

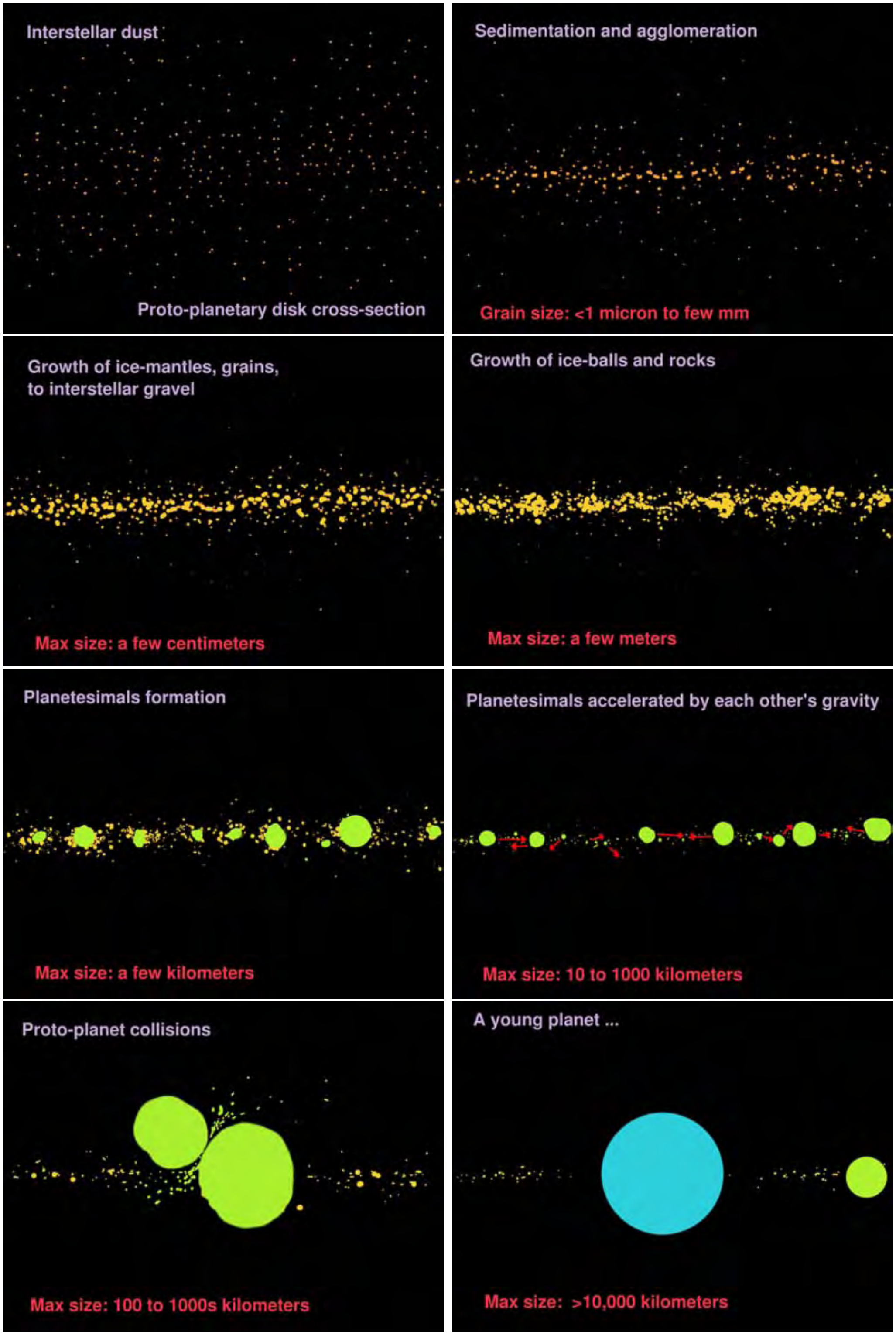


Figure 2.8: Grain and Planetesimal Growth

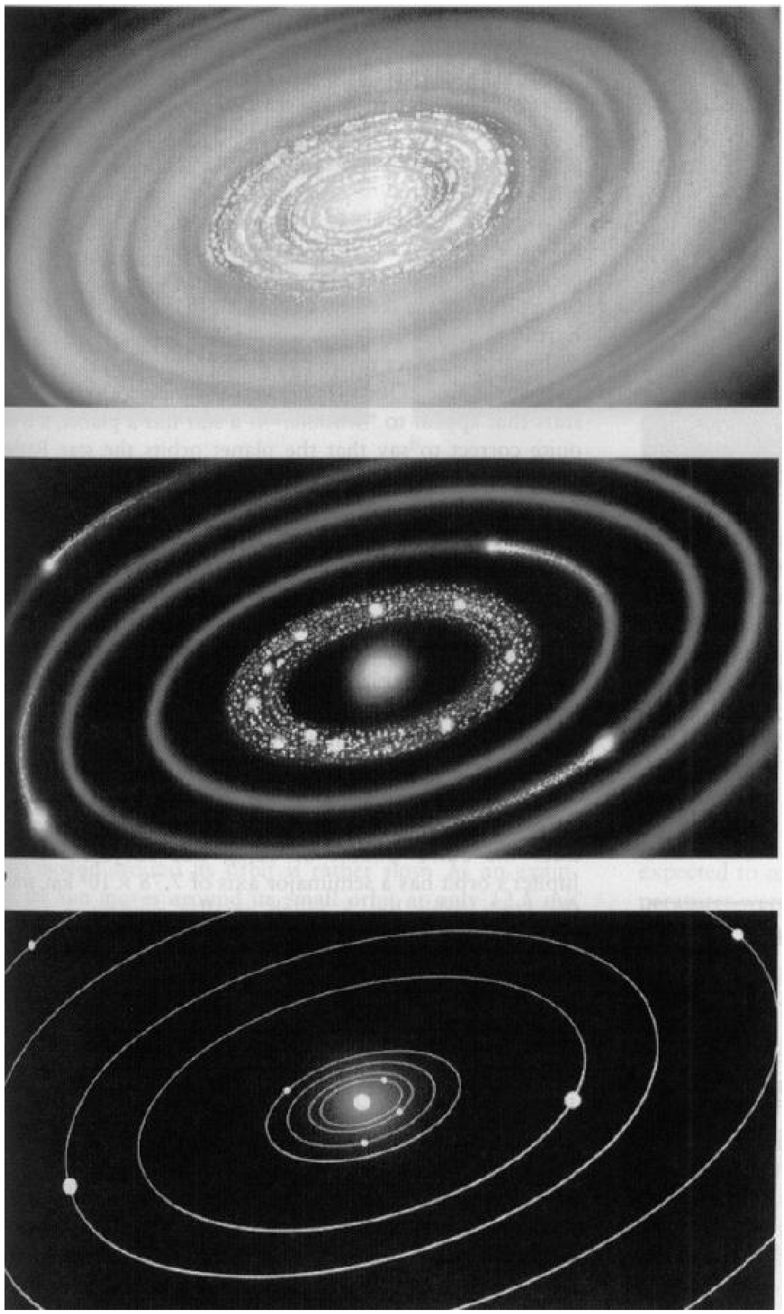


Figure 2.9: Accretion in the Solar System
Over time, the dust, gas, and larger debris in the Solar System is swept up and accreted by the planets. This accretion is still ongoing today in the constant rain of meteorites that land on the Earth, as well as the occasional big impact from larger objects.

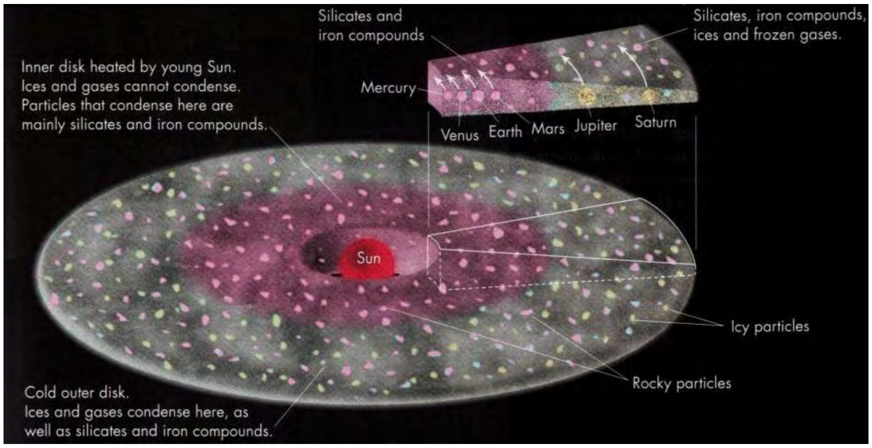


Figure 2.10: Condensation Sequence
Heat from the early Sun prevented volatile ices and gases from condensing in the inner Solar System. The inner planets are therefore composed mainly of rock and metallic materials. Only in the outer Solar System do we find planets rich in gases.

material, than and ices and light gases.

As the molecular compounds freeze out of the solar nebula, the material in the solar material start to *agglomerate* and form sedimentation, falling toward the middle of the disk around the proto-Sun. The particles get larger, growing from microns in size (a millionth of meter) to millimeters across. As particles collide and stick together, they grow larger, into the gravel phase (where typical sizes are centimeters across), to rocks (meters in size), and finally to planetesimals (kilometers across). Once we are at the planetesimals stage, gravity between the largest objects starts becoming important, and the overall process becomes more chaotic since planetesimals are not simply just orbiting the Sun, but are gravitationally interacting with each other as well. This leads to the building of proto-planets hundreds of kilometers across. Collisions at this scale can be extremely violent, with impacts between proto-planetary cores as likely to disperse and eject material back into the solar nebula as they are to continue the accretion.

2.3.1 Cratering

Over time, the amount of material in the solar nebula decreased as it was swept away and accreted onto the planets. Although the absolute amount of debris has dropped during the last 4.5 billion years, it has never been cleared out entirely. Thus we see evidence of multiple **impacts** of material onto planetary surfaces over the history of the Solar System up until today, resulting in the creation of **craters**. The cratering rate at different times in the past has been estimated from observations of and samples brought back from our Moon. Using these estimated rates, and counting the total number of craters on a planetary surface, we can estimate the age of the surface: older surfaces will have had more cratering impacts than younger, newer surfaces.

2.4 The Structure of the Earth

Collisional impacts between bodies hundreds or thousands of kilometers across release enormous amounts of energy, often enough to melt substantial amounts of the material involved in the impact. Even after most of the solar nebula particulate matter and debris had been swept up and accreted onto planetary sized bodies, the planetary cores remained hot and molten. In addition, radioactive materials decayed, which also generating heat. The only way for this heat to escape was for it to conduct to the surface and radiate away into space (see §2.4.2). This process was slow enough that for hundreds of millions of years (and in the case of the Earth, all the way to today), the interiors of the planets remained molten, while the surface cooled enough to form a solid crust.

2.4.1 Crust, Mantle, and Core

When a planet is so hot that the interior is filled with magma, the molten material underwent **differentiation**. Heavy elements like iron and nickel sunk to the center forming a **core**. Lighter elements or elements that could bind together to form less dense minerals tended to find its way further from the core, either in the surface **crust** or the **mantle** in between.

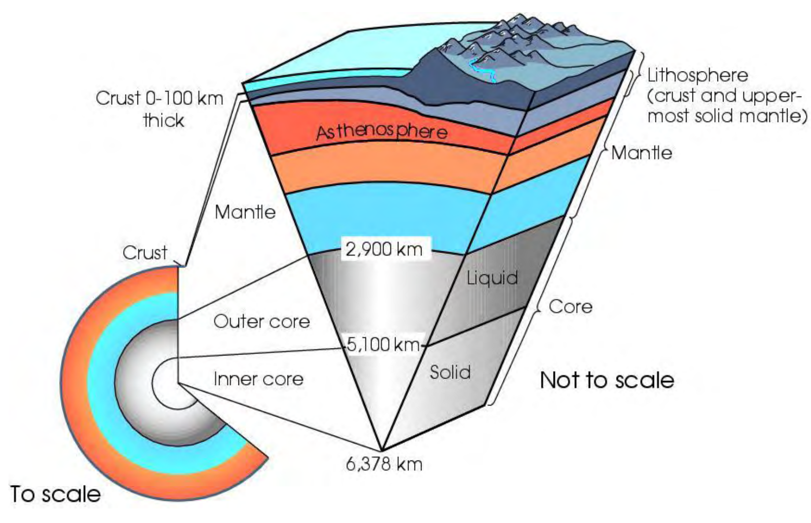


Figure 2.11: Cutaway of the Earth

From seismological studies, and the concerted work of thousands of geochemists and geophysicists, we now have a rough idea of how the elemental composition changes inside the Earth. At the center is dense, solid **inner core**, with a radius of 1200 kilometers. Its composition is mostly iron and about 4% nickel. Surrounding this is a liquid **outer core**, 2290 kilometers thick, and also predominantly iron and nickel, but is less dense, and hence contains some other lighter elements, such as oxygen, sulfur, and potassium.

Beyond the core is the **mantle**, a region with high temperatures and still under enormous pressure. The mantle is composed mostly of minerals containing oxygen (O), silicon (Si), calcium (Ca), iron (Fe), and magnesium (Mg), with smaller amounts of aluminum (Al), sodium (Na), and chromium (Ch).

Finally the outermost layer is the **crust**, about 5 kilometers thick underneath the oceans, and about 30 kilometers thick under the continents. The crust consists of lighter minerals than those found in the mantle, and includes **basalt** and the even less dense **granite**. The continents in fact consist mostly of granitic rocks that float atop the denser basalt in the lower crust and upper mantle.

2.4.2 Plate Tectonics

Because the center of the Earth is much hotter than the surface, the heat will naturally try to escape. There are three ways by which heat energy can be transported:

1. **Conduction:** Heat is transported through a material, or a physical medium, spreading from the hotter to cooler regions. *Example: The heat from your stove is* conducted *through your pot, then* conducted *again through water to cook your food.*

2. **Convection:** Heat causes a material to expand, growing less dense. It rises through a medium, and then dissipates its heat at the surface or at the top of the medium. Once it cools, it grows denser, and then sinks back down. *Example: The waxy blobs inside a lava lamp. The material gets heated at the base of the lamp; it falls back down after cooling at the top of the lamp.*

3. **Radiation:** Heat is transported away in the form of electromagnetic radiation. No physical medium is necessary for the radiation to escape. *Example: the reddish-orange glow emanating from the spiral metal burners on your electric range. You can feel the heat from them even if your hand is a several inches away from the burner.*

Heat can escape from a planet via all three of these methods. Once heat energy reaches the surface, it **radiates** away into space. Much of the heat from the center of the Earth is *conducted* out to the surface, by diffusing through solid (or liquid) rock. In parts of the mantle, the physical conditions are just right that *convection* can occur. Note that at this region of the Earth, the mantle is still completely solid. However convection can occur for material in all three phases of matter (gas, liquid,

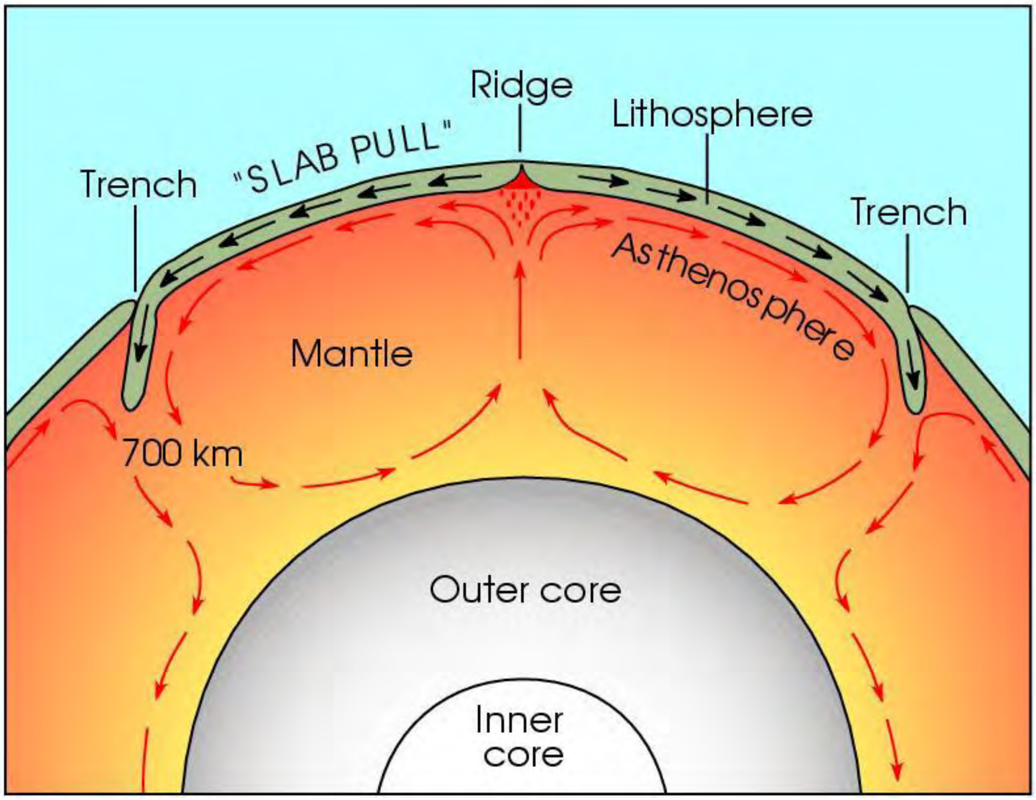


Figure 2.12: Convection in the Mantle System

and solid). There is thus a slow, but inexorable transport of rising, hot material through the mantle, up toward the crust.

As the planet cools, a rigid layer forms on the surface. This is the **lithosphere**, which on the Earth is about 100 kilometers deep, and which contains the crust as well as the upper mantle. Below this is the **asthenosphere**, a layer between 100–350 kilometers below the surface, where the pressure, temperature, and chemical compositions of the minerals result in a plastic, semi-molten, sluggishly flowing material. The convection currents in the asthenosphere bring up mantle material upward, where they create "hot spots" in the Earth's crust.

The mantle material also pushes up against the lithosphere, and drags on it, creating stresses and cracking it into large pieces. These continent-sized **plates** are pushed along by the asthenosphere, in a process called **plate tectonics**. Magma from the mantle can **upwell** to the surface, and create new surface crust, pushing apart existing plates. The Mid-Atlantic Ridge is one of the **mid-ocean ridges**, where new material is brought up to the surface, and pushes aside slightly older oceanic crust, resulting in **sea-floor spreading**. Conversely in other places on the Earth's surface, plates are pushed underneath or **subducted** below other plates. This

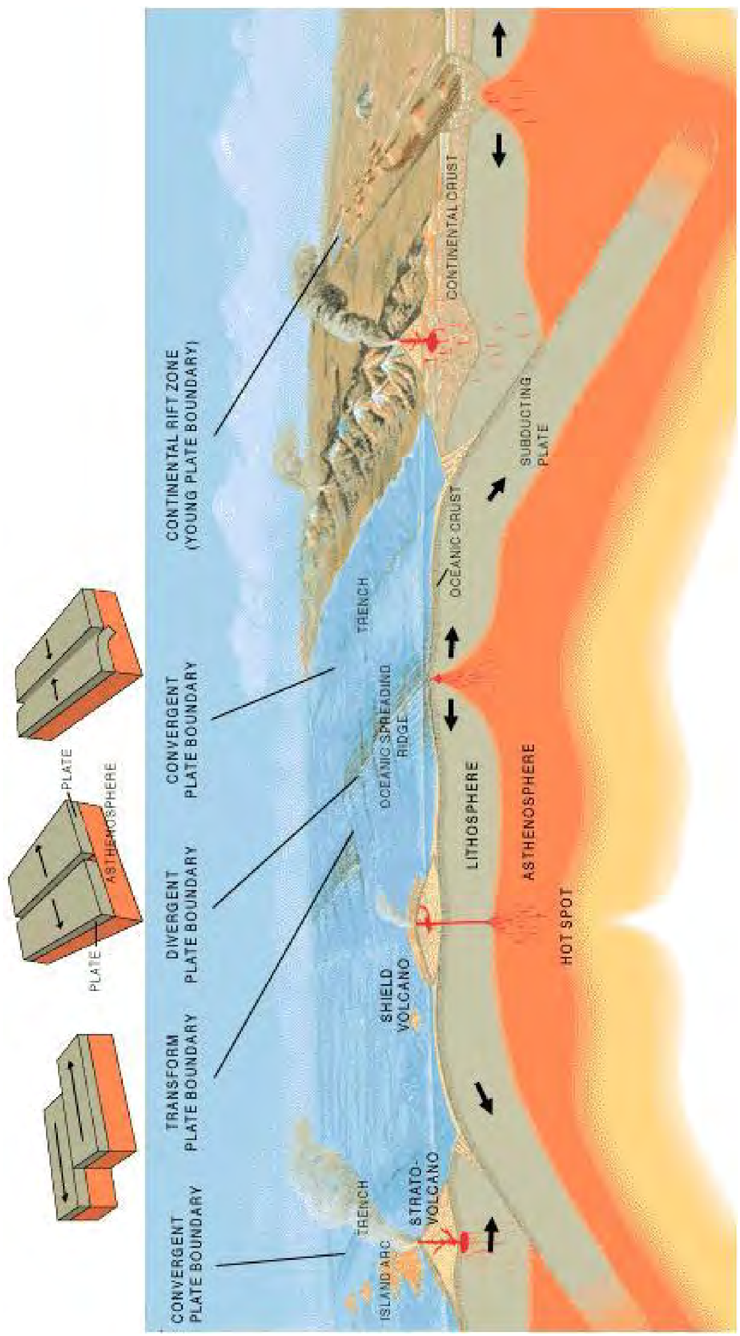


Figure 2.13: Plate Motions

As the mantle pushes against the lithosphere, rigid plates spread apart as new crust is created in oceanic ridges and continental rift zones. A plate converging against another can be subducted underneath the crust where it is swallowed up by the mantle. Finally, a transform boundary is found between two plates moving in parallel past each other.

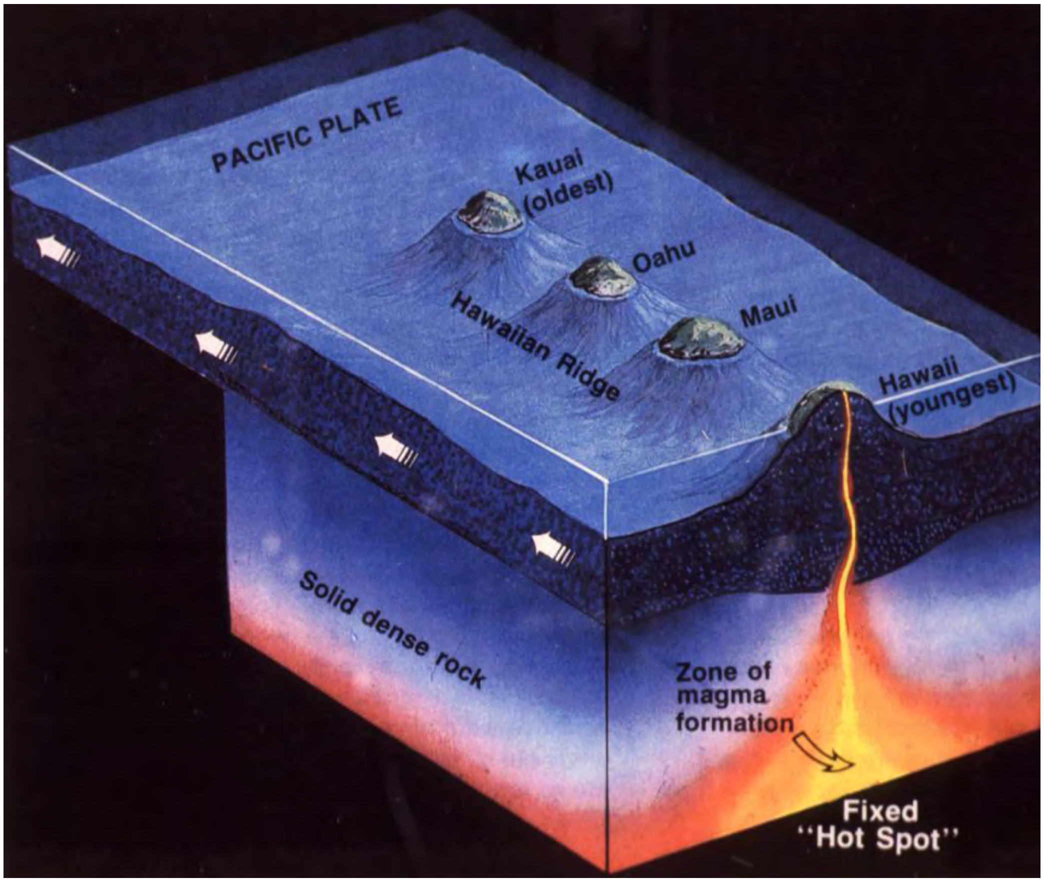


Figure 2.14: Hawaiian Islands
The Hawaiian Islands were formed as the Pacific Plate moved over a fixed "hot spot." The youngest island is still over the hot spot, and as a result, is still growing.

cold, older material is carried down underneath the continental crust back into the asthenosphere, where it is recycled. The convection cycle thus starts with hot magma rising through the mantle to form new crust, and ends with cold, older crust destroyed elsewhere on the surface. Because of this **plate recycling**, it can be difficult to find rocks on the Earth's surface that are as old as the planet.

It is in the plate boundary regions where **volcanoes** are prevalent. Mt. St. Helens in Washington is a volcano associated with the Pacific plate subducting underneath the North American plate. Iceland is a volcanic island created from the spreading of the mid-Atlantic Ridge. Volcanoes can also form in the middle of plates, such as the Hawaiian Islands near the center of the Pacific plate which is moving over a mantle "hot spot." Over time, the eruption of lava leads to the formation of a new volcanic island. As the plate moves the island past the center of the hotspot, a new upwelling of material would leads to another island forming further down the chain. The most

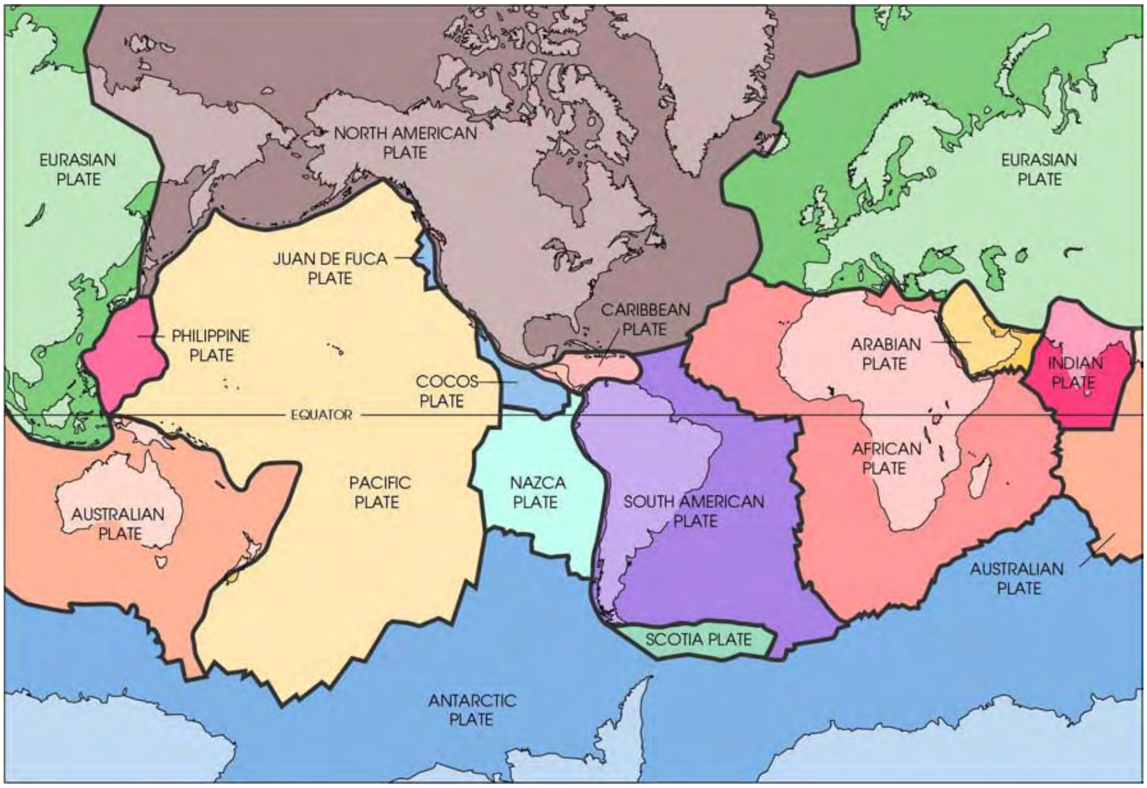


Figure 2.15: Global Tectonic Plates

recent one is the "big" island of Hawaii, less than half a million years old. The islands further to the northwest in the chain (Maui, Oahu, and Kauai) are increasingly older, with Kauai 5.1 million years old. Islands which formed earlier are also smaller today because they have had more time to erode away.

As a plate subducts under another plate, it can crumple and push up crustal material. The Indian plate is doing just that as it drives north into the Eurasian plate, creating the massive Himalayan mountains. A similar type of activity is occurring in the western hemisphere where the Nazca plate impacting South America is creating the Andes mountain range.

Finally where two plates are sliding parallel to each other's boundaries, fractures on the surface called **faults** are created. When the two plates are stuck and cannot slide, they build up stress until they become unstuck. This sudden release of energy results in earthquakes.

The Earth's plates move at speeds of centimeters per year or less. Extrapolating backward in time tens to hundreds of millions of years ago, the configuration of the continents must have been very different. Over time, the continents can be driven together to form a single **supercontinent**. This last happened about 290 million years ago, when all of the continental land mass was in Pangaea. This supercontinent

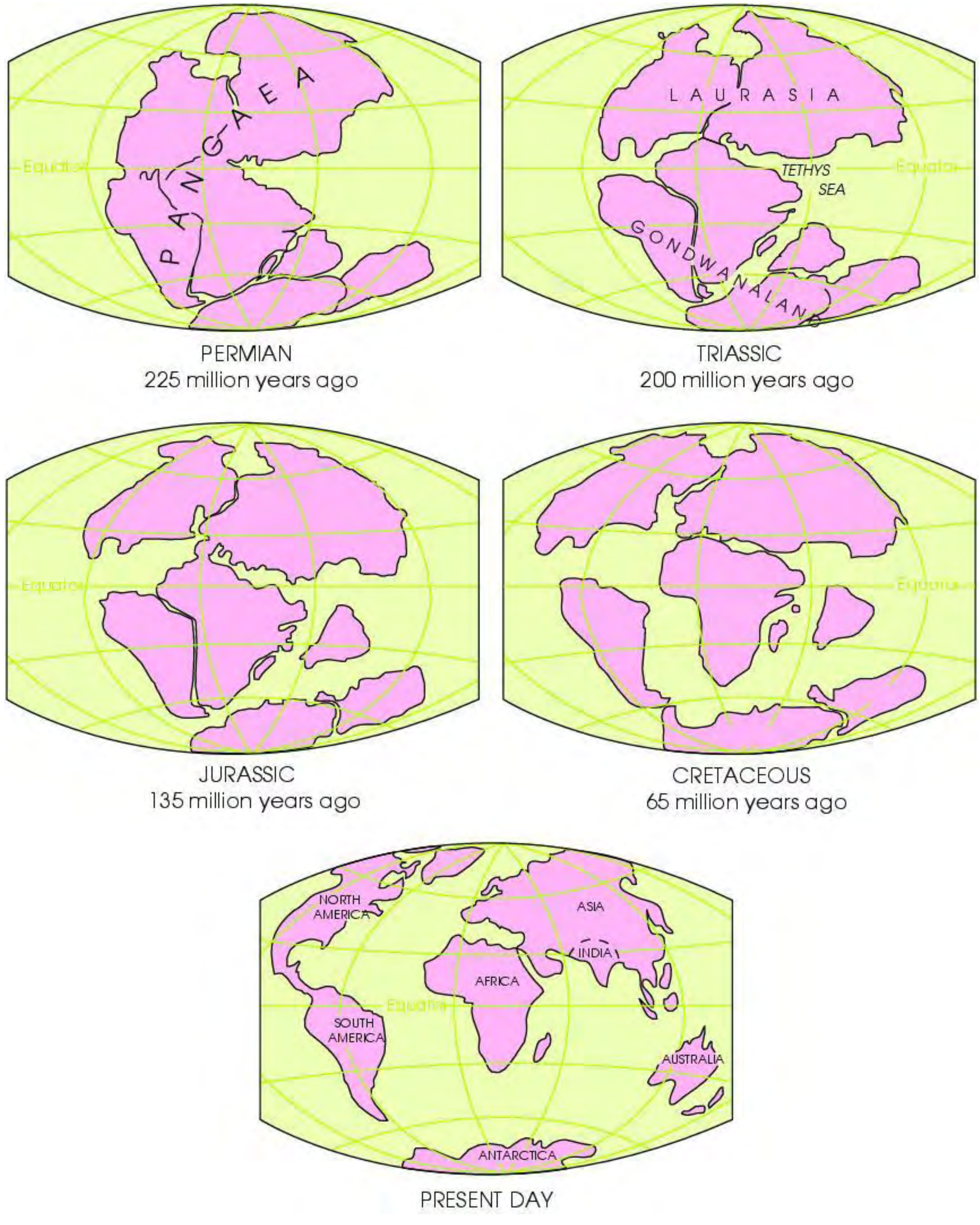


Figure 2.16: The Break Up of a Super Continent
Although the plates move extremely slowly, given enough time, the crust of the Earth will appreciably change in appearance. 200 million years before the present, all of the continents were inter-connected in the single land mass Pangaea.

subsequently split up into two large landmasses, Gondwanaland and Laurasia, which themselves eventually broke up into the continents that we are familiar with today.

2.5 Comparative Planetology

The Earth is geologically alive because it still has enough heat in its core and mantle. This heat creates the active convecting asthenosphere, which results in geological activity that alters the surface. New crustal material is created and destroyed. Furthermore energy from the Sun drives the wind and water on the Earth further weather and erode the surface.

As the Earth gradually loses its interior heat, we would expect its geologic activity to decrease. The lithosphere will thicken, and the top of the convecting asthenosphere will get deeper. At some point in the distant future, surface tectonic and volcanic activity will cease once there is not enough residual heat to drive convection. The planet will be effectively "dead."

We do not have to wait the billions of years necessary to know what a geologically dead surface looks like. Because the four (five if you include the Moon) terrestrial planets have different sizes and compositions, they started out with different amounts of primordial heat and heating from radioactive decay. The smaller bodies tended to radiate away their heat energy faster, and so are all thought to be geologically dead. Let us examine them one by one.

2.5.1 The Moon

Our best current theory on the origins of the Moon states that it occurred during the violent, chaotic collisional phase of planetary accretion. The Moon has a composition similar to the Earth's mantle, and hence appears to be depleted in metals such as iron and nickel. Although the proportion of refractory materials are similar to the Earth's mantle, it does contain anomalously fewer volatile elements, such as aluminum and titanium, which have higher boiling points. This suggests that lunar material might have been severely heated in the past, which would have driven off the volatiles. Finally the proportions of isotopes of oxygen on lunar rocks is exactly the same as the Earth's, suggesting that it formed roughly in the same place in the solar nebula.

These curious facts about the composition of the Moon has lead to the **giant impact hypothesis** for its formation. When interplanetary collisions between bodies hundreds and thousands of kilometers across were prevalent, it is thought that the young Earth collided with a Mars size body, approximately 4.5 billion years ago. Simulations show that if the collision was oblique, hitting the Earth at a slight angle as opposed to being head-on, the resulting impact "splashed" out material from the Earth's mantle as well as disintegrating the striking body. This collision formed a debris ring in orbit around the Earth, consisting of mantle material with most of the volatiles boiled off. This material eventually coalesced into the Moon.

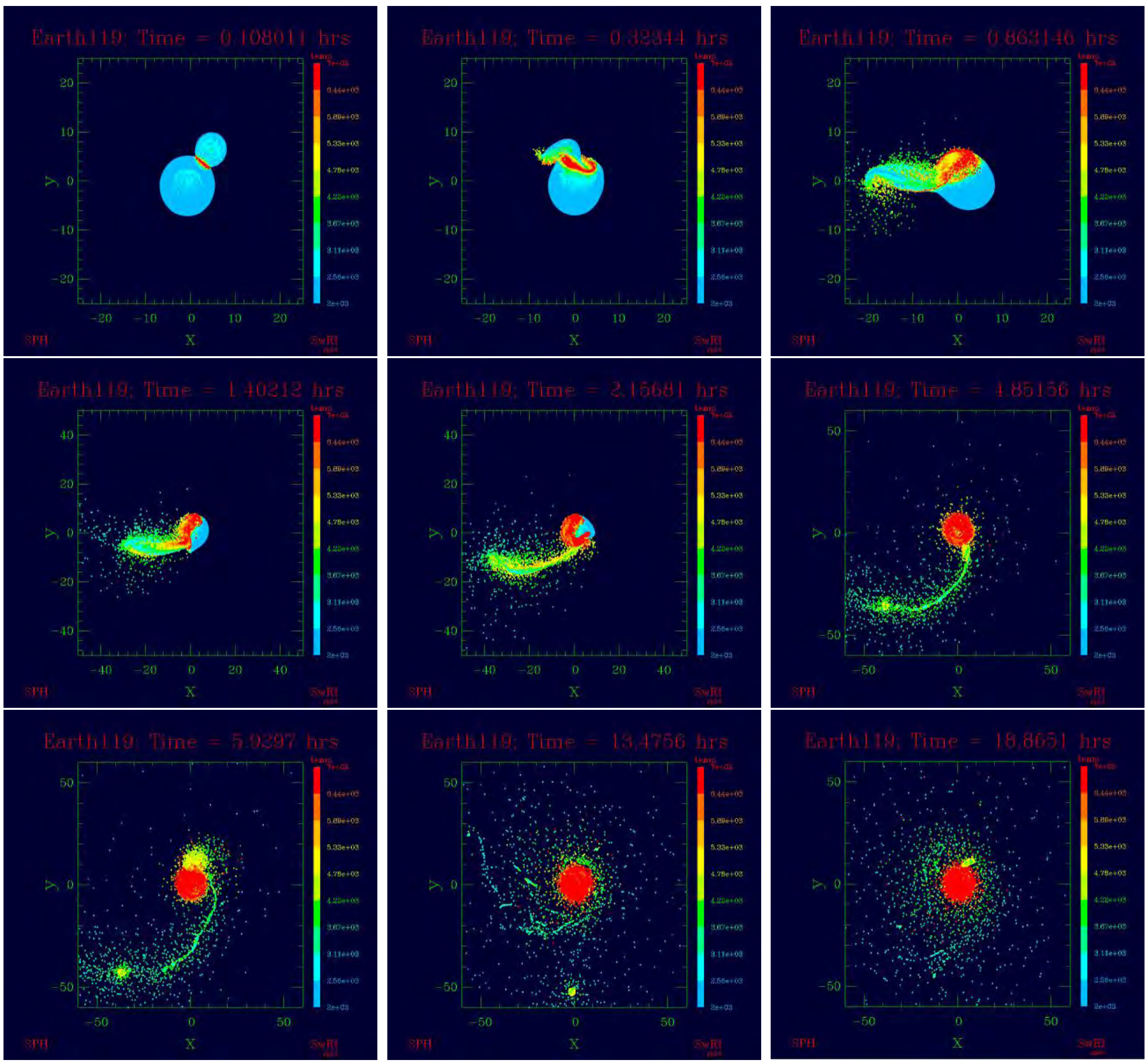


Figure 2.17: Formation of the Moon

Snapshots from Robin Canup's simulation of the impact on the Earth that ejected debris into orbit which eventually coalesced into the Moon.

Because the Moon is a much smaller body than the Earth, it is thought to have lost most of its primordial, internal heat long ago. However surface features on the Moon do give indications of its molten past, where **magma oceans** laid just underneath the surface crust. The dark flat **maria** or "lunar seas" are a hint of the last of the ancient basaltic lava flows that upwelled to the surface after spectacular surface impacts. The current maria that we see today were created between 3.2–3.85 billion years ago. However in the 500 million years preceding this period, there were earlier episodes that created many of the giant **impact basins** on the lunar surface.

Since the end of the early bombardment period, the Moon has had time to cool off, enough to the point where the lithosphere is thought to be 1000 kilometers thick. Volcanism died out on the Moon about 3 billion years ago. The surface is geologically dead, with no continental plates, or lighter basaltic rock atop granitic crust. In the last 3 billion years, the only activity that has changed or broken up its surface has been from the steadily declining pelting of meteorites that create new craters and churn up the surface to create the lunar **regolith** soil.

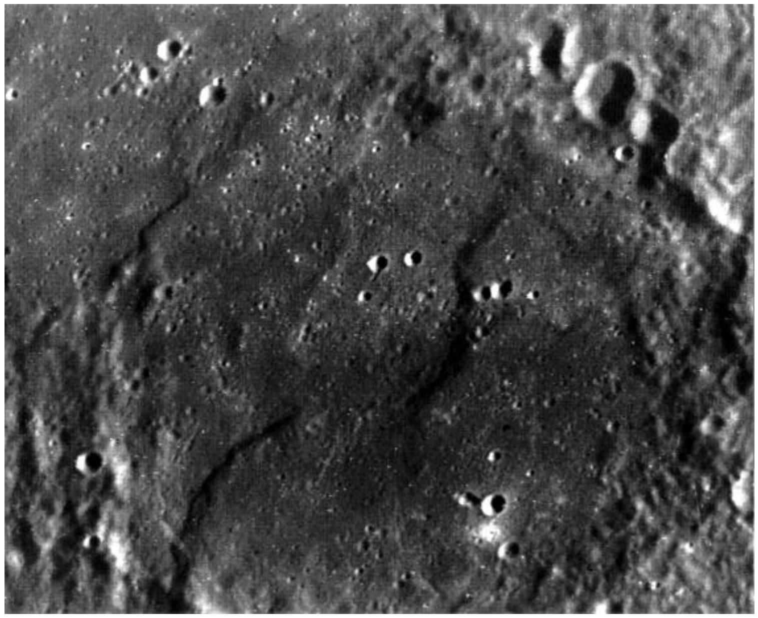
2.5.2 Mercury

Mercury is close in appearance to our Moon, with a surface that is marked by craters, impact basins, and lava flows. The surface has small ridges which snake hundreds of kilometers across ancient lava plains and impact regions. These are thought to be the result of compressional forces in the lithosphere as Mercury cooled. Mercury also has a weak magnetic field (about 1% that of the Earth) which suggests an abnormally large iron core compared to the other terrestrial planets. Some have suggested that this is an indication of giant impact event, which removed much of Mercury's mantle.

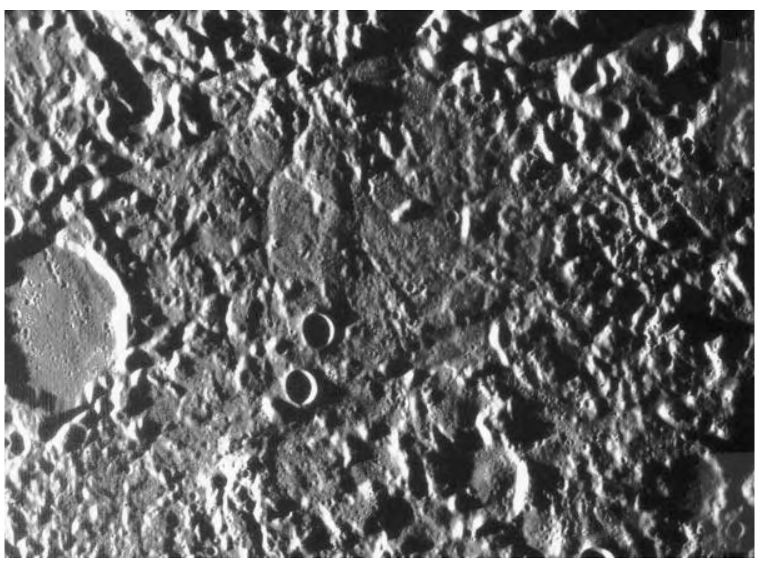
Mercury's small size meant that it lost most of its internal heat long ago, and has been geologically dead for most of the history of the Solar System.

2.5.3 Venus

Venus is similar to the Earth in size and density, so one might expect that the core and mantle of this sister world would be similar as well. However the images derived from the Magellan radar mapping campaign of the planet has revealed that Venus does not appear to have crustal plates. It does have indications of past (although no present) volcanic activity. A massive catastrophic overflow of magma which covered most of the surface occurred about 500 million years ago, which wiped out most of the impact craters older than this date. The best theory for why this has happened is the following. Although convection currents in Venus' mantle are responsible for volcanism just as it is on the Earth, the lithosphere might be too thick for heat to be lost continuously. Instead the internal heat builds up, trapped underneath the upper mantle, until it reaches to the point where the mantle rock partially melts. The less dense mantle material rapidly "overturns." The resulting escape of the trapped



(a)



(b)

Figure 2.18: The Surface of Mercury

Images from the Mariner 10 spacecraft: (a) Scarps (cliffs) on the surface formed possibly by compressional forces; (b) "jumbled" terrain in another part of the surface thought to be the result of tectonic activity.

magma leads to a massive, catastrophic resurfacing of the planet. After this heat is released, the lithosphere cools again, trapping the internal heat, and the cycle begins anew.

2.5.4 Mars

Mars has about 1/10th the mass of the Earth, but its density makes it most similar to the Moon. The Martian lithosphere does not appear to be broken up into tectonic plates that move over time. As a result, Olympus Mons, which is 24 kilometers in height, may have been able to grow into the largest volcano in the Solar System because it has been sitting over the same hot spot for a very long time. Another fact that lends support to this idea is that Olympus Mons is located in the Tharsis "bulge," a region 4000 kilometers across that is elevated compared to the rest of the planet, suggesting a mantle plume is pushing up against the crust. Valle Marineris, the giant canyon system on Mars, might be the result of the cracking of the crust as the Tharsis region was uplifted.

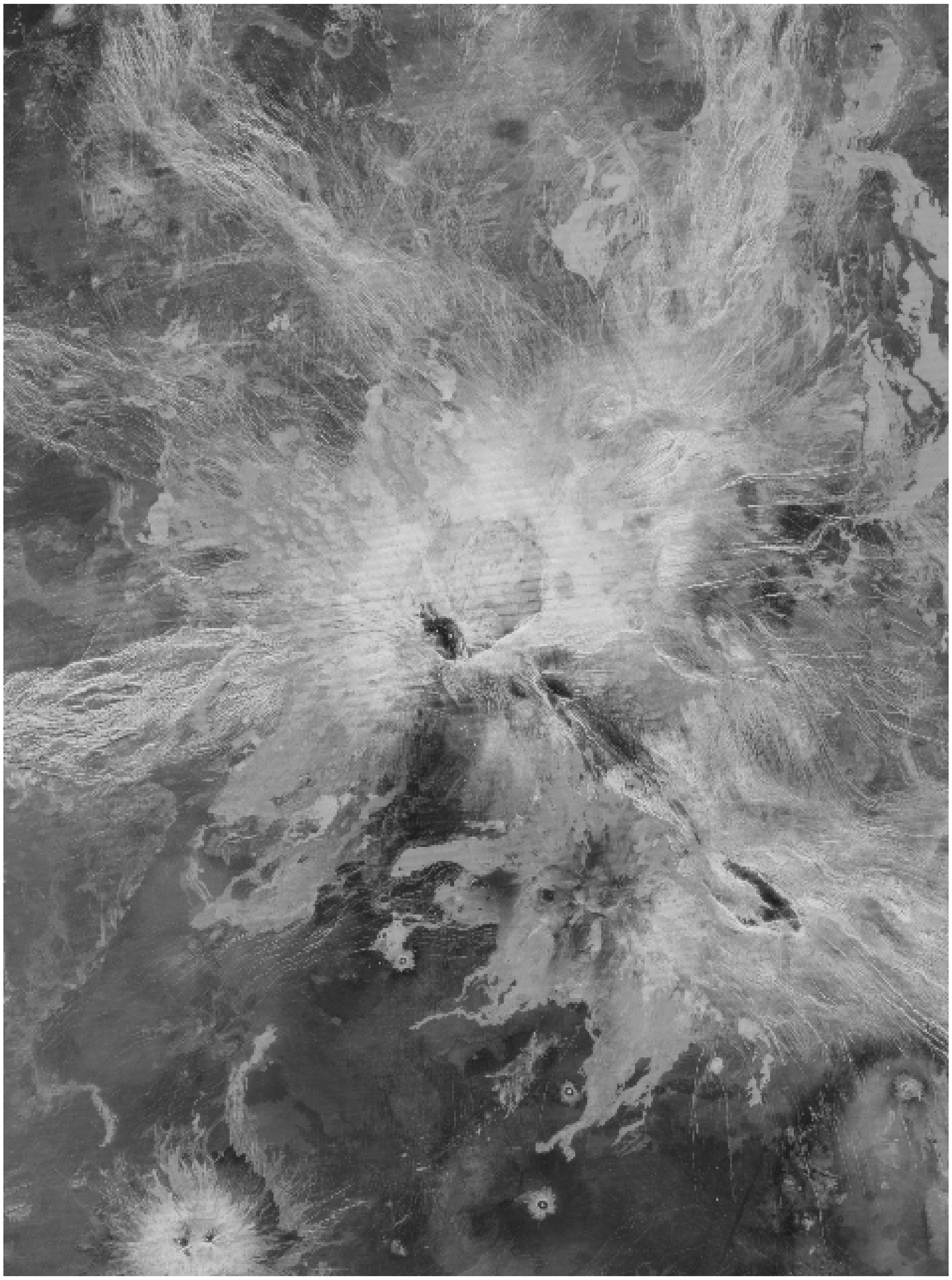
2.6 Terrestrial Atmospheres

The three larger terrestrial planets—Venus, Earth, and Mars—have substantial atmospheres. Their compositions seem to be very different at first glance. Carbon dioxide (CO_2) is a major constituent on both Venus and Mars, while it remains a trace gas on the Earth. Molecular nitrogen (N_2) makes up more than three-quarters of the Earth's atmosphere, but is about 3% on the other two planets. The Earth also contains a substantial amount of oxygen (O_2) that is not found in substantial quantities on either Mars or Venus. Venus also contains anomalous constituents such as sulfur dioxide (SO_2), and clouds made up of tiny droplets of sulfuric acid (H_2SO_4) instead of water. The total amount of atmosphere, or equivalently the atmospheric pressure on the three planet surfaces, are also radically different.

2.6.1 Venus

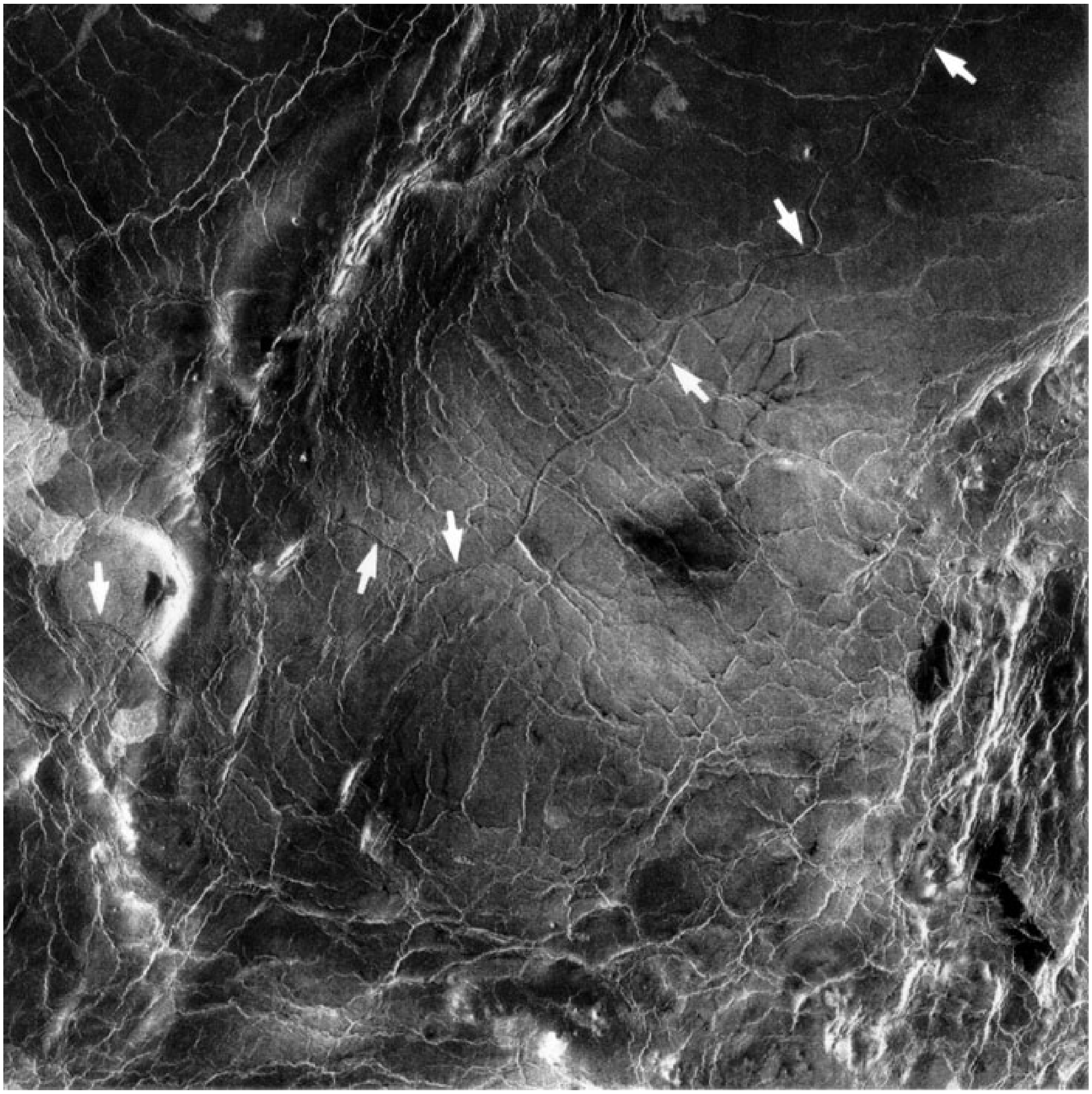
The atmosphere of Venus is the thickest out of all of the terrestrial planets. The surface pressure is 90 times that of the surface pressure on the Earth. Instead of 14.7 pounds/square inch of force pressing down on the Earth, one would experience 1320 pounds/square inch, roughly equivalent to the water pressure 3000 feet beneath the sea. Even more startling is the surface temperature of 480° C (890° F), which is constant over the entire surface regardless of whether it is day or night.

The reason for the extraordinarily hot planetary surface is due to the **greenhouse effect.** Venus has a bright, sulfuric acid cloud deck located about 60 kilometers above the surface of the planet, which reflect about 77% of the incoming sunlight back out into space (which explains why Venus appears so brilliantly in our evening



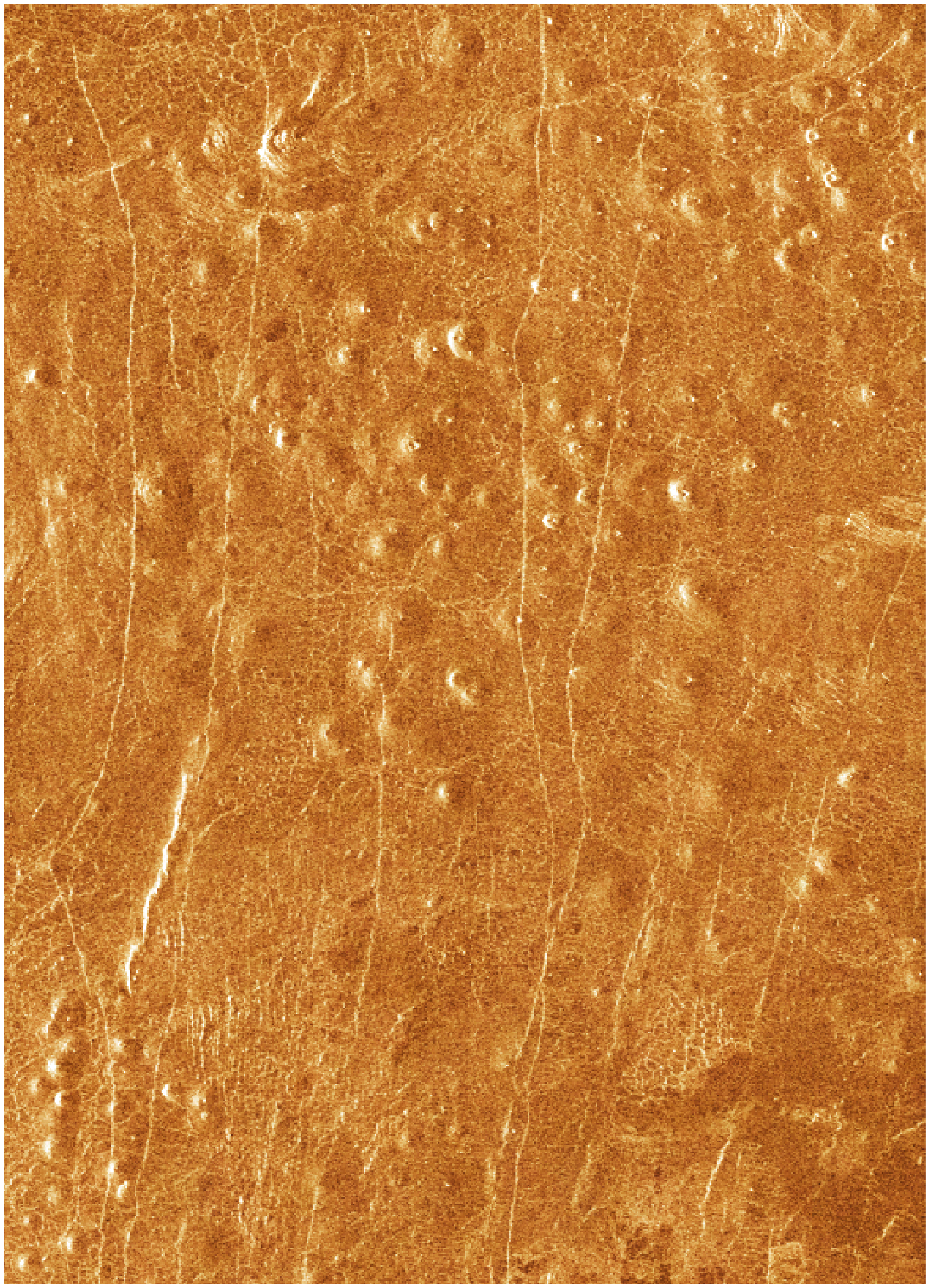
(a) Venusian Volcanoes

Figure 2.19: *Volcanoes on the eastern plains of the Aphrodite continent; shown are Ozza Mons (center) and Maat Mons (lower right). Extensive bright lava flows can be seen emanating from the two volcanic calderas. Fine lines criss-crossing the image are the result of tectonic activity in the crust after the lava had solidified.*



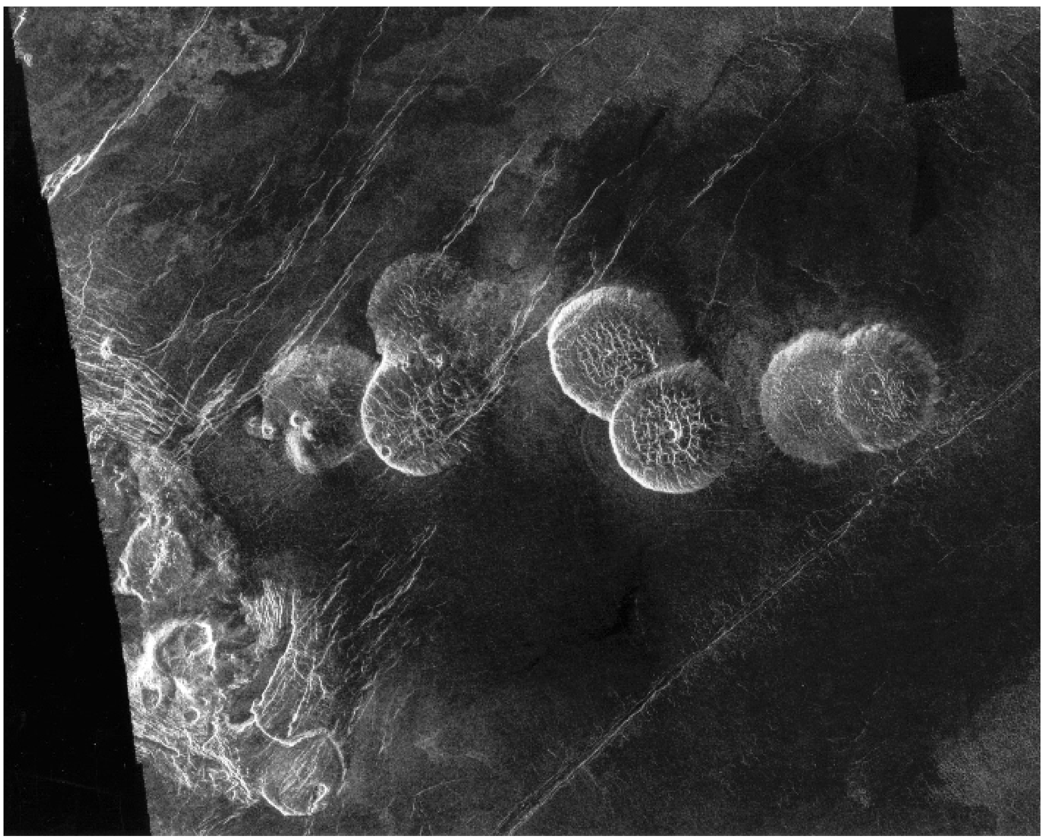
(b) Baltis Vallis

Figure 2.19: (Cont.) *The longest river in the Solar System is Baltis Vallis, a lava channel that is more than 4200 miles in length. This image shows a 360 mile segment of it meandering through Atalanta Planitia on Venus.*



(c) Venusian Shield Volcanoes

Figure 2.19: (Cont.) *A field of dozens of shield volcanoes, similar to Mauna Loa in Hawaii. Each volcano is roughly two miles in diameter.*



(d) Pancakes

Figure 2.19: (Cont.) *"Pancake" domes in Alpha Regio. These volcanic features are about 15 miles across, and at least 2000 feet high. They are thought to have formed from especially viscous lava that flowed only a short distance before congealing.*

(e) Tessera

Figure 2.19: (Cont.) *An area of "chaos" terrain or* tessera *in Alpha Regio. The image is roughly 80 miles across, and shows a complex set of faults and ridges created from multiple waves of tectonic activity originating in many different directions.*

Figure 2.20: Valle Marineris

The largest canyon system in the Solar System is located on Mars. Valle Marineris is 2500 miles long, reaches depths of up to 4 miles, and at its widest, is 125 miles across. The canyon is probably a giant tectonic crack in the crust, formed by the rising of the Tharsis Bulge.

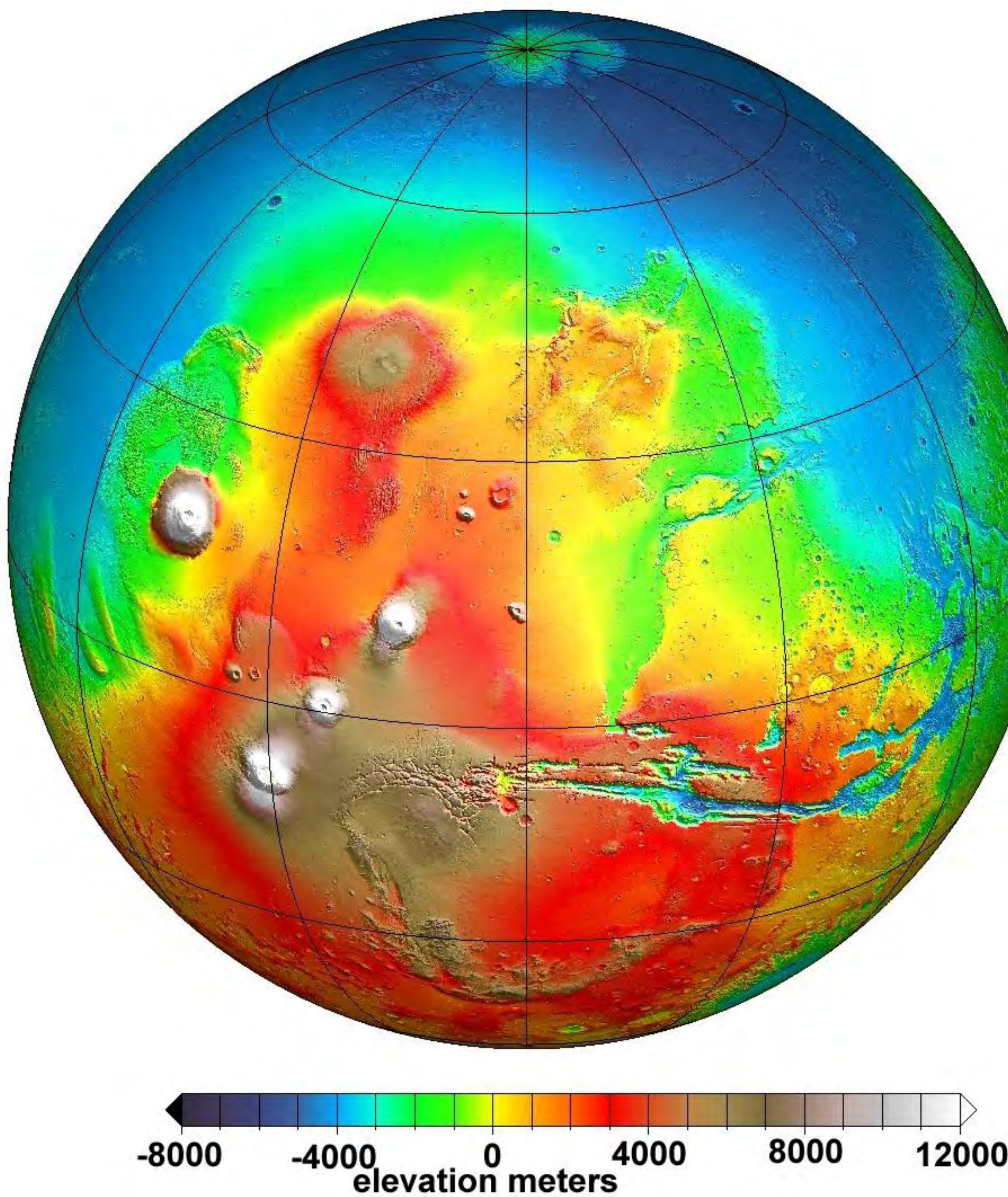


Figure 2.21: The Tharsis Bulge

Mars Orbiter Laser Altimeter data, which shows the elevation change over the surface of the planet. The entire Tharsis region is uplifted and color coded red, showing a altitude of roughly 4000 meters. Also sitting in Tharsis is the largest volcano in the Solar System, Olympus Mons (white peak on the left side), as well as three smaller volcanoes (from top right to bottom left): Ascraeus Mons, Pavonis Mons, and Arsia Mons.

or morning sky). The fraction of the sunlight that filters through to the surface strikes the ground and heats it; the surface rock can only radiate to cool. However the thick CO_2 atmosphere is not transparent to heat (or **infrared**) radiation. The CO_2 molecules absorb the heat and reradiate it back, effectively trapping the heat in the atmosphere. The greenhouse effect on Venus keeps it hundreds of degrees hotter than if there was no atmosphere on the planet.

2.6.2 Earth

Early in its history, the Earth's primordial atmosphere consisted mostly of CO_2, just like Venus, and water H_2O. It is thought that much of the Earth's atmosphere resulted from **outgassing** from volcanoes. However because either Earth formed with more water, or it lost less water into space because it was further from the Sun, much of the CO_2 became dissolved in Earth's oceans. A chemical reaction resulted in the formation of carbonic acid. This weak acid reacted with minerals on the sea floor leading to carbonate rocks. Much of the original CO_2 atmosphere has therefore been locked up in the Earth's rocks. After the evolution of plants, photosynthetic life began to generate oxygen, leading to an oxygen-, rather than a CO_2-dominated atmosphere roughly 2–2.5 billion years ago.

Oxygen is a very reactive gas; O_2 will bond with metals to form rust, burn with hydrogen to form water, and react with many other minerals and compounds to form oxides. If all of the photosynthetic life on the Earth were to be wiped out, the oxygen would disappear on a timescale of tens to hundreds of millions of years.

Having an atmosphere with 23% O_2 also gives rise to a substantial amount of **ozone** (O_3) on the Earth. This gas is concentrated in the Earth's **stratosphere**, between 10–50 kilometers, forming an **ozone layer**. Ozone is the only gas in the Earth's atmosphere that significantly absorbs ultraviolet radiation from the Sun, and does so quite efficiently despite its low concentration. The harshest of the Sun's UV radiation can wreak havoc on the DNA of living organisms. Therefore it has been theorized by some that the invasion of the continental land masses by plant and later, animal life, could not begin until the ozone shield was in place.

The trace CO_2 provides a slight greenhouse effect on the Earth.

2.6.3 Mars

The Martian atmosphere is very thin, less than about 1% that of the Earth (or the equivalent of about 100,000 feet altitude here on Earth). The current atmosphere is mostly CO_2, and is similar to Venus in some its trace constituents. However planetary scientists have discovered evidence that the Martian atmosphere was anywhere from 10–00 times thicker in the past. Studies of heavy gases, such as argon, suggests that substantially more atmosphere was outgassed by Martian volcanoes than exist there on the planet today. The amount of water ejected in volcanic eruptions would have been enough to form a water layer anywhere from tens to hundreds of meters thick over the entire surface. Furthermore, orbital images of the surface show features that seemed to have been carved by running water on the surface, which would have required a much denser atmosphere.

How did Mars lose its atmosphere? Lighter gases would not have been held on as strongly by Mars' weak gravity, and would have leaked off into space. Much of the CO_2 could also have been locked up as carbonate rocks on the Martian surface. Oxygen atoms liberated from CO_2 and H_2O molecules by UV radiation from the Sun would quickly oxidize with minerals. Without the greenhouse effect from the substantial CO_2 atmosphere, the planet slowly cooled and dried up, with the remaining water freezing into the permafrost and the polar ice caps.

Table 2.3: Composition of Terrestrial Atmospheres

Venus		Venus		Venus	
CO_2	96.5%	N_2	78.1%	CO_2	95.3%
N_2	3.5%	O_2	20.9%	N_2	2.7%
SO_2	0.015%	H_2O	$< 4\%$	Ar	1.6%
H_2O	0.010%	Ar	0.93%	O_2	0.13%
Ar	0.007%	CO_2	0.034%	CO	0.07%
H_2	$< 0.0025\%$	Ne	0.0018%	H_2O	0.03%
CO	0.002%				
		Pressure at Surface			
	90 atm		1 atm		0.006 atm

Lecture 3

The Outer Solar System

The Outer Solar System is dominated by the gas giants, Jupiter, Saturn, Uranus, and Neptune. As we learned in §2.3, the present planets originated from the solar nebula, with the elements solidifying according to a **condensation sequence**, where hardy **refractory** compounds are able to survive the warmer temperatures close to the Sun, while the delicate **volatile** compounds can solidify only far from the Sun. The inner planets are rocky, with thin atmospheres. The outer planets however grow in a region of the nebula full of icy volatiles. The rich abundance of frozen gases can be observed in the moons of the outer Solar system which tend to be mostly icy bodies, as well as the large host of small bodies beyond the gas giants, including the ninth planet, Pluto, the growing list of objects in the **Kuiper Belt**, and the **comets** in the **Oort Cloud.**

Let us first concentrate on the giant gas planets, which we have quite a bit of information on, because of the fleet of spacecraft which have flown by them. Probably the most noticeable differences between the gas planets and the terrestrial planets are their size, and hence why they are referred to as giant planets. Jupiter has a diameter 10× that of the Earth's; Saturn is slightly less than 10× across; Uranus and Neptune have diameters about 4× that of the Earth.

The second important fact is the difference in densities: the outer Solar System giants have much lower densities. Because the terrestrial planets are mostly rock, with few volatiles, their total densities are similar to that of rock: between 3.9–5.5 grams per cubic centimeter. The gas planets have densities between 0.7–1.6 grams per cubic centimeter. Water has a density of 1 gram per cubic centimeter, so an oft heard analogy is, that if you could find an ocean large enough, you would be able to "float" Saturn, the least dense planet. In addition to their voluminous gaseous atmospheres, the outer planets do appear to consist largely of volatile compounds, like ices, as well as low-density liquids such as liquid hydrogen. The chemical compounds that solidified and condensed into the proto-planets in the outer Solar system include frozen compounds of hydrogen, such as water (H_2O), methane (CH_4), and ammonia (NH_3). Thus in addition to the rocky materials in the solar nebula, the outer planets *formed from both rock and ice*, roughly in equal 50-50 proportions.

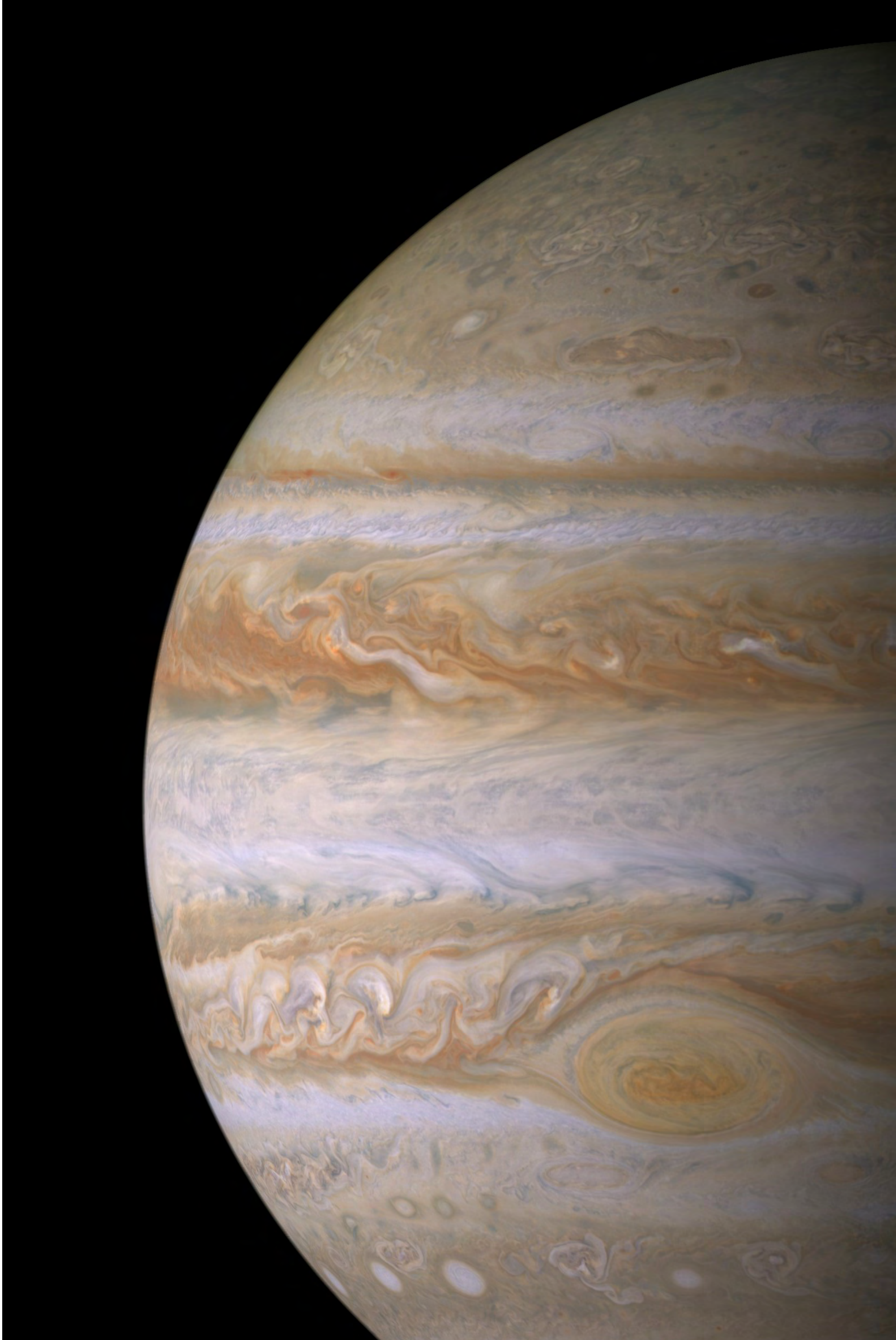
Also significant is that the outer proto-planet cores were able to grow to such a size that they were able to gravitationally accrete the surrounding hydrogen gas. The more gas they accreted, the stronger their gravitational fields, which lead to an increased accumulation. This **runaway accretion** resulted in the sweeping up of nearly all of the nebular gas in the vicinity of the giant planets. These gas giants, especially Jupiter, became enormous. The inner terrestrial planets were too small, and hence their gravitational fields were too weak to hold onto these light gases.

3.1 Jupiter

The largest of the gas giant planets is Jupiter, fifth planet from the Sun, and appropriately enough, named after the king of the Roman gods. Jupiter is so large in fact that its mass is roughly $2\,{}^{1}/_{2}$ times greater than all of the rest of the other planets combined. It contains about 71% of all of the planetary mass, and currently also has the most extensive moon system in the entire Solar System, with 61 known natural satellites. (However with the Cassini spacecraft in orbit around Saturn, the number of moons around the ringed sixth planet, will bound to increase with new discoveries past its current stable of 29 known objects.)

When we look at Jupiter, we are not seeing a "surface" of the planet, but are looking only at the top of the cloud-decks of this world (and this is true of the other gas planets as well). The atmosphere consists mostly of hydrogen, with helium comprising almost a quarter of the mass. This ratio of hydrogen to helium is very similar to the Sun, which lends support to the solar nebula origin of the planets as well as our primary star. The clouds of Jupiter are broken up into bands of bright **zones** and darker **belts**, running parallel to the equator. The zones have features that are white-ish or yellow-ish, while the belts have gray-brown and reddish-tinges. The colors in these atmospheric bands come from compounds that happen to be just minor atmospheric constituents. Many of these compounds are suspended as ice crystals of ammonia, water, and ammonium hydrosulfide in the upper atmosphere. They are suspended in the same way that water droplets on Earth can condense and form clouds that float in the Earth's nitrogen-oxygen atmosphere.

The clouds on Jupiter also show many turbulent patterns, similar to what we see from weather satellites imaging the water clouds on Earth. High-speed and high-altitude winds on Jupiter (and on Saturn), similar to the jet stream on Earth, blow at speeds up to 150 meters/second (or 338 mph). Because the gas giants have no solid surface, it is difficult to estimate the rotation speeds of the planets as a whole. However based on the radio emissions from electrical storms, the rotation of Jupiter is $10^{\mathrm{h}}55.5^{\mathrm{m}}$. When we observe the rotational rates of the belts and zones in Jupiter's atmosphere, we find that they rotate at slightly different speeds, with the equatorial bands moving faster around the planet, giving a mean rotation period of $9^{\mathrm{h}}50^{\mathrm{m}}$.



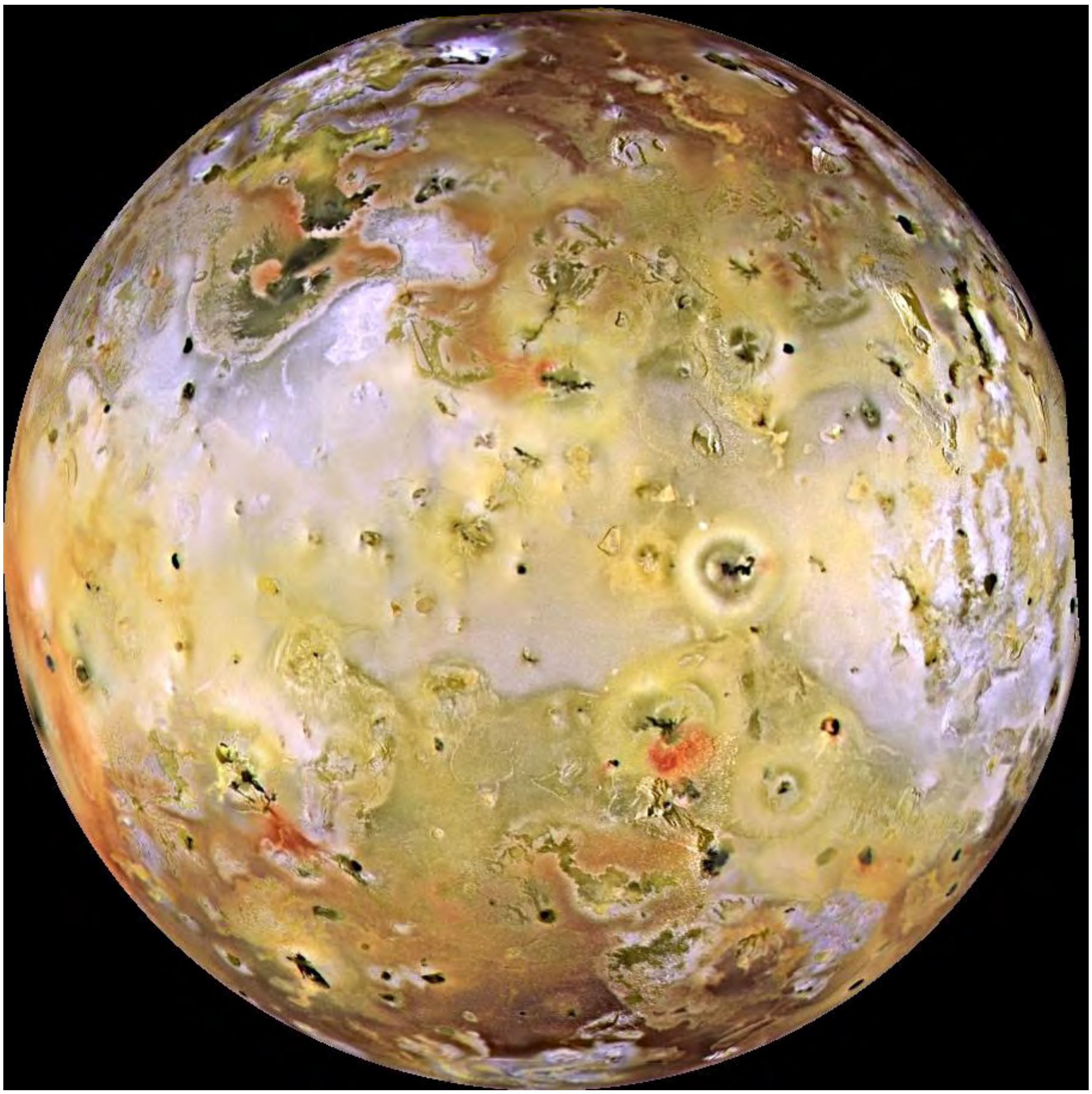


Figure 3.2: Io

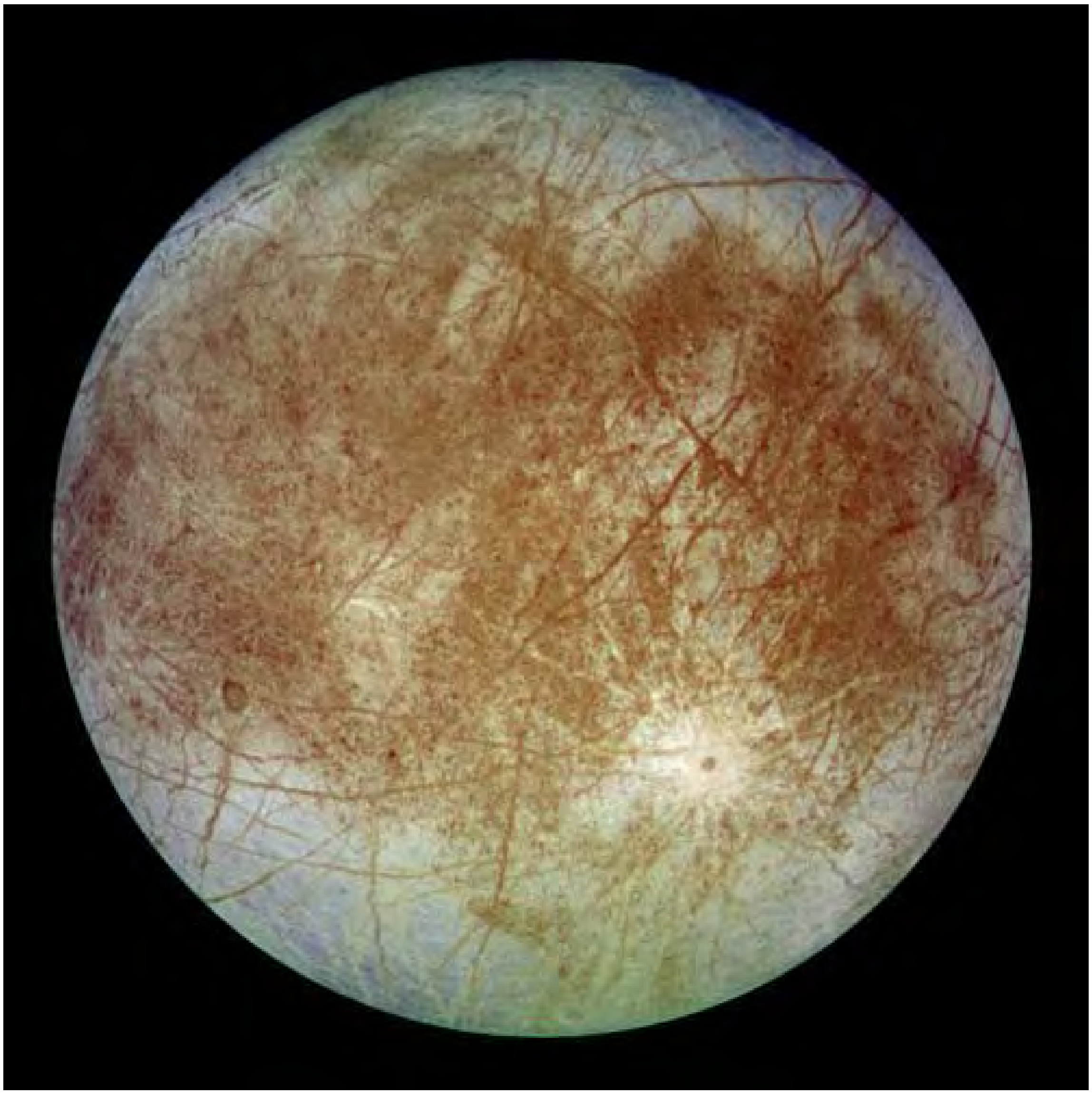


Figure 3.3: Europa

Figure 3.4: Ganymede

Figure 3.5: Callisto

Jupiter also has the **Great Red Spot**, a feature that has been in the atmosphere of Jupiter for at least 300 years, ever since the first report of it by the Italian astronomer G. D. Cassini in 1665. The Great Red Spot appears to be the largest of hurricane-like Jovian storms. Smaller storms are white-ish and transient, and have been observed to get swallowed up by the Great Red Spot. Although the Great Red Spot might look small in a telescope, it is actually about three times the diameter of the Earth.

3.1.1 The Galilean Moons

In the early 1970s, most astronomers had thought they had the outer Solar System figured out. Because of their parent planets' distances from the Sun, the moons of the giant planets would be uninteresting, frozen ice-balls, with their surfaces cluttered with craters. However the astronomers were proven spectacularly wrong when these moons were imaged by the Voyager spacecraft which traveled through the outer Solar System starting in 1979. The images beamed back by our robotic explorers show an incredible variety of surfaces and features, and worlds with characteristics that no one had imagined could have existed.

In orbit around Jupiter are four of the largest moons in the Solar System. They were first seen when Galileo pointed his telescope at Jupiter and observed them in orbit around the gas giant, giving support to the Copernican Revolution, by suggesting that not everything in the universe moved around the Earth (see §2.2). Hence they are known as the **Galilean moons** today. The Galilean moons are, in order of their distance from Jupiter, Io, Europa, Ganymede, and Callisto (all named after mythical companions to the god Jupiter).

The innermost **Io** is perhaps the most dramatically different world in the whole Jovian system. Its surface is a patchwork of bright colors, yellows, tans, oranges, blacks, and whites. In addition to the Earth, Io is the only other known volcanically active body in the Solar System, with multiple eruptions that were recorded by the Voyager (and later the Galileo) spacecraft. These volcanoes spew plumes of material 100 km above the surface of Io. Many of the round, crater-like spots on the surface are indeed, not craters, but **calderas**, the points at the tops of volcanoes where material erupt from. In fact Io contains very few visible impact craters. The volcanoes constantly resurface the moon, covering over new impacts. Io is even more geologically active than the Earth, if one considers that its entire surface can dramatically change on a timescale of a decade or two. Most of the ices and other light volatile materials have been driven off of Io, by the volcanic activity. The moon thus probably has a composition more similar to the terrestrial planets than it does to its fellow Galilean moons. The volcanoes have been found to eject sulfur compounds, which is black when molten, but can cool to a variety of yellow and reddish-orange colors. The black areas around the volcanic calderas are thus thought to be pools of molten sulfur lava. The white regions are probably sulfur dioxide condensates.

The second furthest Galilean moon is **Europa**, which from the Voyager images appear as the smoothest object in the entire Solar System. Europa is exceptionally featureless, with again, very few impact craters, mountains, or any obvious deep canyons. The surface is creamy white in color, composed mostly of water ice. However higher resolution images show that it is criss-crossed by numerous faint lines, which are shallow grooves and cracks on the surface of the ice. The lack of impact craters tells us that Europa's surface is relatively young. Theoretical modeling of the moon suggest that underneath the surface ice (which can be 100 km deep) of Europa is a liquid water ocean. Images from the Galileo spacecraft show that some of the surface ice has been broken up into enormous iceberg blocks kilometers across, and re-frozen in the recent past (within the last 100 million years). Europa's liquid water ocean is probably responsible for the resurfacing of the moon. Impacts and other tectonic activity can cause water to leak up to the surface, where it can refreeze and cover over impact features.

The next moon out is **Ganymede** which with a diameter of 5262 km, is the largest moon in the Solar System, larger than our Moon, and bigger than the planets Mercury and Pluto. Ganymede has two types of terrain, a darker surface with many craters, suggesting it is older, and a lighter colored terrain that appears grooved in high resolution images. The darker crust thus appears to have split open, allowing new icy material to erupt up and flow to the surface. Ganymede might even have had early plate tectonics, with the surface ice split into blocks that floated atop a more plastic ice mantle. Although Ganymede is probably too cold to have full fledged motion of the surface lithosphere plates like that found on the Earth, it might be a transitional case between completely frozen ice-balls, and warmer, geologically active worlds.

The last of the Galilean moons is **Callisto**, which is the third largest moon in the Solar System, larger than our Moon, but just slightly smaller than the planet Mercury. Callisto is heavily cratered, and shows very little evidence of tectonic activity, in either fracturing or ice flows. Callisto does not seemed to have been altered at all by any internal heating, and hence surface-wise, is the most primordial of all of the Galilean moons.

3.1.2 Tidal Heating

What accounts for this strange range of characteristics between the four Galilean moons? There is a clear correlation between the amount of internal heat and geological activity of each moon, and its distance from Jupiter. The innermost Io is the most active, with its constant volcanic activity, and the amount of activity decreases outward: Europa has a liquid ocean and surface ice blocks that break apart and rearrange themselves before refreezing; Ganymede has grooved ice flows; Callisto has the least activity of the four.

All four of the Galilean moons happen to be **tidally locked** to Jupiter. Like our

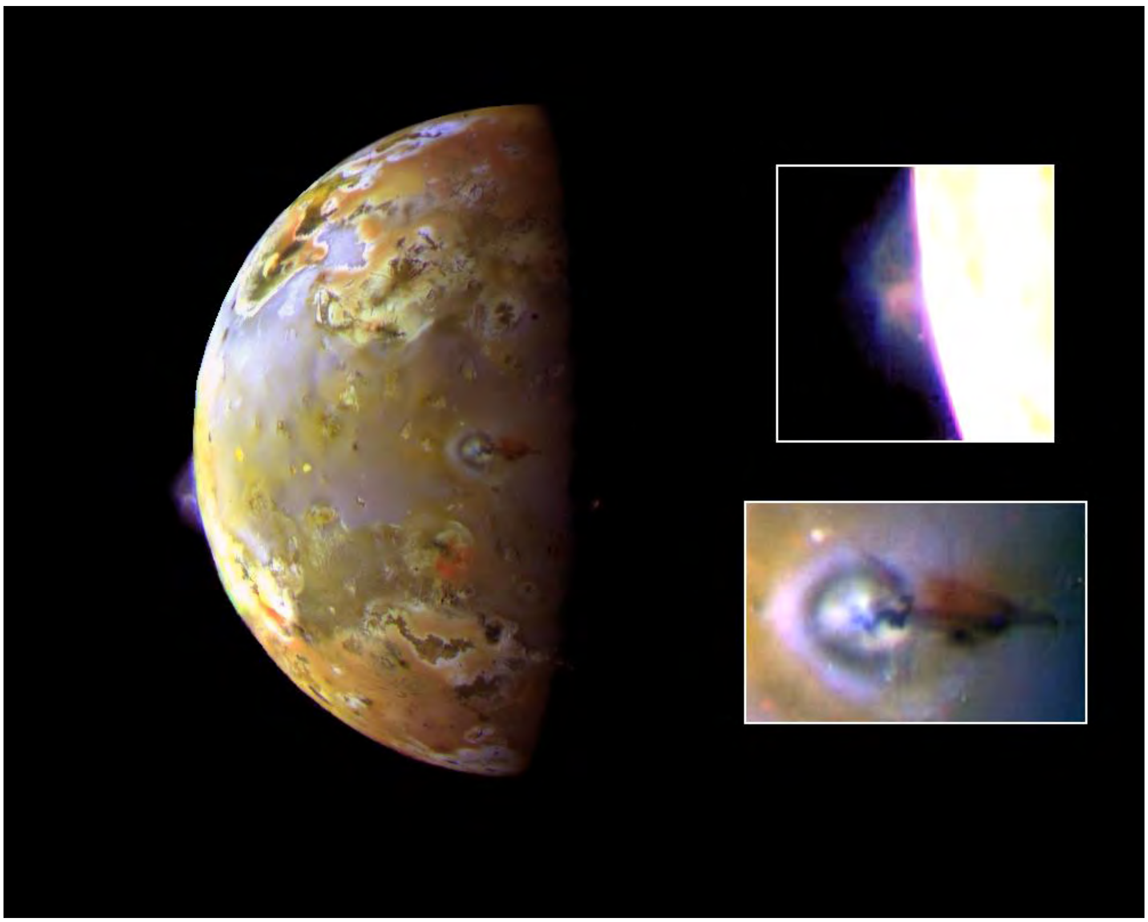


Figure 3.6: Volcanoes on Io

Io as imaged by the Galileo spacecraft, showing several active volcanoes including Pillan Patera (left image, and top right inset) and Prometheus (bottom right inset).

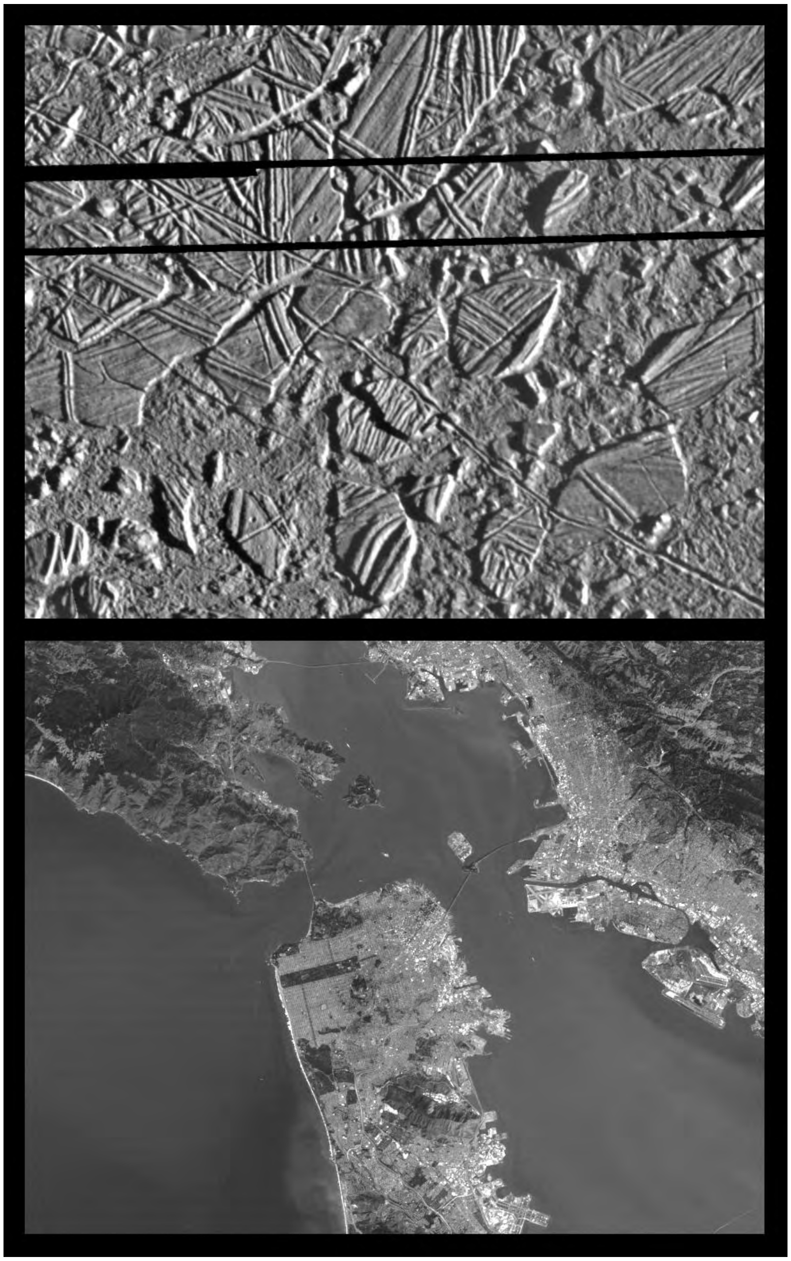


Figure 3.7: Ice Rafts on Europa

One piece of evidence of a liquid water ocean beneath the surface of Europa are these blocks of ice on the surface, which have apparently broken up, floated about before rearranging themselves and re-freezing. Using the fine linear ridges on these blocks and employing a jigsaw puzzle-solving strategy, one can even reconstruct what the original surface must have looked like before it got jumbled. The top image is at the same scale as the bottom one, which shows the San Francisco Bay region; the size of the area displayed is 34 kilometers by 42 kilometers. The smallest features visible in both images are 54 meters across.

Moon's tidal locking to the Earth, the Galilean moons show the same faces to Jupiter. If we were to land on the surface, Jupiter would always appear to be in the same spot in the sky. We can turn on **meridian** lines in Cosmic Atlas for each moon. These are lines representing the 0° meridian line (equivalent to the Greenwich meridian line on the Earth) extended out into space. For tidally looked moons, the 0° meridian line is *defined* to be directly below the planet. Thus we can see that as each moon orbits the planet, the meridian line extended out into space will always be pointed at the planet.

The four moons are also in **orbital resonance** with each other. If we were to move forward through time, and watch the orbits of Io and Europa, we would find that Io orbits twice around Jupiter for every single orbit of Europa. We say that Io and Europa have a **2:1** resonance. Similarly when we compare Io and Ganymede, Io orbits four times for each of Ganymede's single orbit, so there is a **4:1** orbital resonance between the two. Finally Callisto and Io show a **9:1** resonance.

Io's orbit around Jupiter is not perfectly circular, but is slightly elliptical (again because of the gravitational tugs from the other Galilean moons). The force of gravity from a body decreases in strength with distance. As a result the side of Io facing Jupiter will be tugged on harder than the side facing away from the planet. This *differential* effect of gravity is known as a **tidal force**. The proximity of Io to Jupiter accentuates the tides, which stretch and squeeze the moon each time it moves through its orbit.

The orbital resonances between Europa, Ganymede, and Callisto on Io means that Io gets regular gravitational tugs from each of these moons. The strength of this tidal tug is proportional to the moon's distance from Io, so that Europa has the strongest (and most frequent) effect, while the gravitational pull from Ganymede and Callisto are less. The cumulative effect is that Io gets flexed and pulled at least once in its orbit by Jupiter, and perhaps more than once if it passes the other Galilean moons. This constant stretching heats the interior of the moon until it turns mostly molten.

The tidal effects on Europa from Jupiter is less, because it is further away from the planet. However the cumulative effect including that of the other moons means that Europa is warm enough to have a liquid water ocean and significant tectonic activity on its surface.

Ganymede is even further out, has much weaker gravitational interactions with Jupiter and the other moons, giving it far weaker tectonic activity. Finally Callisto might be so distant that its inner heat is not being replenished; it may have turned into an example of the canonic ice-ball that astronomers had originally predicted would fill the outer Solar System before the 1970s.

3.2 Saturn

Saturn is similar to Jupiter in having a yellow-ish orange appearance. Saturn's atmosphere also has the same main constituents, hydrogen and helium, but also many of the same minor elements such as ammonia and sulfide compounds, as Jupiter's. However its cloud structure is not as detailed or turbulent as Jupiter's. This is attributed to the fact that Saturn is further away than Jupiter, almost twice as far, and hence receiving a quarter as much light and heat from the Sun. Jupiter also has retained much more of its primordial heat of formation (so much so that it radiates more internal heat than it receives from the Sun). Jupiter's weather systems are thus driven as much as from heat originating deep in its core as it does from sunlight. Although Saturn does not have large hurricane-like storms, similar to Jupiter's Great Red Spot, it does have bright zones and darker bands.

Saturn is best known for its spectacular **ring system**, which were first seen by Galileo, who in 1610, reported fuzzy blobs on either side of the planet. He thought Saturn was a triple planet, with two smaller spheres to either side of the main disk. However in 1655, Christian Huygens with a better telescope, realized that the blobs were in fact parts of a ring system encircling the planet.

The rings are not flat sheets of solid material, but consist of billions of ice particles with individual orbits about Saturn. Telescope observations from Earth were able to distinguish a handful of main rings separated by gaps, the most prominent of which is the **Cassini division** which is visible even in telescopes used by amateur astronomers. However the Voyager spacecraft revealed that the rings consisted of thousands of fainter ringlets. Even the gaps contained darker rings. The ring particles themselves range in size from dust-like fragments, to ice rocks, to giant blocks the size of automobiles and houses. Although the ring system from the inner edge to the outer edge stretches 274,000 km, dynamical forces probably keep the rings very thin, probably 100 m thick or less.

The Voyager observations revealed that tiny moons are located within some of the ring gaps. These are **shepherd moons** and they appear to keep the ring particles restricted to certain zones by gravitational resonances similar to the orbital resonances of the Galilean satellites.

What are the origins of Saturn's ring system? Computer simulations show that the rings are not stable over the lifetime of the Solar System. They would dissipate and be lost to Saturn over as short a time period as several hundred million to a billion years. The only way to provide enough material to the current ring system is if an icy moon around Saturn broke apart because of tidal forces from the planet. Because the material in Saturn's rings will eventually spiral into the planet, we might be just lucky enough to be living at a point in time after the precursor moon has broken up, perhaps only a few hundred million years ago, but before most of its material has been lost.

Saturn's mid-sized moons are icy bodies which show mild tectonic activity. Objects like **Mimas**, **Enceladus**, **Tethys**, **Rhea**, and **Dione** are heavily cratered, but also show grooved surfaces that hint at some of the internal heating that could have occurred on these bodies. Mimas is also unusual by the existence of an enormous impact crater, roughly $^1/_4$ the diameter of the moon, evidence of an impact that

Figure 3.8: Rings Close-Up I

A detailed true color look at Saturn's rings from the Cassini spacecraft. Visible is the B ring, the bright sandy band curving from the left to the upper right, which itself is composed of thousands of ringlets.

Figure 3.9: Ring Close-Up II

Yet another Cassini view of the rings, taken in the spacecraft's first orbit about Saturn. The faint, outermost F ring is visible as well as the moon Mimas in the lower right part of the picture.

almost smashed apart the icy body.

Saturn's largest moon **Titan** is the second largest moon in the Solar System (after Ganymede). Titan is especially remarkable because of its thick, dense atmosphere that is 1.6× thicker than that of the Earth. Although it is 90% nitrogen, the other 10% consists of methane (CH_4), ethane, acetylene, ethylene, and other hydrogen compounds. These are chemicals that were probably created by reactions driven by sunlight, and hence are similar in character and origins to smog that you find in polluted urban areas on Earth. Simulations of the atmospheric chemistry of Titan show that methane can play the same role there as water does on the Earth. In addition to being a minor component of Titan's atmosphere, methane could rain out of the clouds as snow or icy hail, and may also exist on the surface in the form of methane oceans. Titan is of interest to the astronomers because its atmospheric and surface chemistry might be similar to that of the primeval Earth. This is why the Cassini spacecraft is planning to drop the Huygens probe down into Titan's atmosphere.

3.3 Uranus and Neptune

In contrast to the orange-red-yellow colors of Saturn and Jupiter, **Uranus** and **Neptune** have distinctly different blue-ish hues, with a slight green-ish tinge. Their atmospheric composition is however mostly hydrogen (H_2) and with a smaller fraction of helium (He). Both Uranus and Neptune appear blue because of methane in the atmosphere, which is in greater proportions than is found on Jupiter or Saturn. The methane preferentially absorbs red and orange light, scattering back into space blue light. Haze in the atmosphere also scatters blue light more than red light, adding to the effect. Uranus has no bright and dark bands, and is in fact nearly featureless. Because of both its distance from the Sun and the rather small amount of internal heat, the atmosphere is not very dynamic at all.

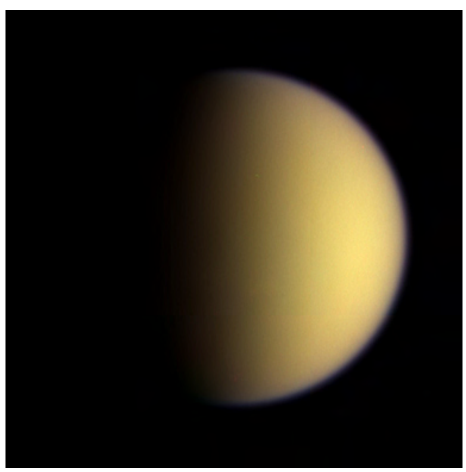
Uranus has an extreme axial tilt, tipped by 98° on its side with respect to its orbital plane. Astronomers speculate that just as a large impactor collided with the Earth to create the Moon, the early proto-Uranus had a collision so powerful that it changed the planet's rotational axis.

Neptune is similar to Uranus in color, but its atmosphere does have significant white-ish cloud features, as well as a "Great Dark Spot." Neptune is even further from the Sun than Uranus, so it must have more primordial heat in its interior to explain its weather features. This could be the result of a greater proportion of decaying radioactive elements, or impacts heating up Neptune's core early in the history of the Solar System.



(a) Mimas (b) Enceladus (c) Tethys


(d) Rhea (e) Dione



(f) Titan

Figure 3.10: Moons of Saturn

Five icy moons around Saturn as seen by the spacecraft (a) Voyager 1, (b) Voyager 2, and (c,d,e) Voyager 1. The true color picture of Titan was taken more recently by the Cassini spacecraft.

Figure 3.11: Uranus

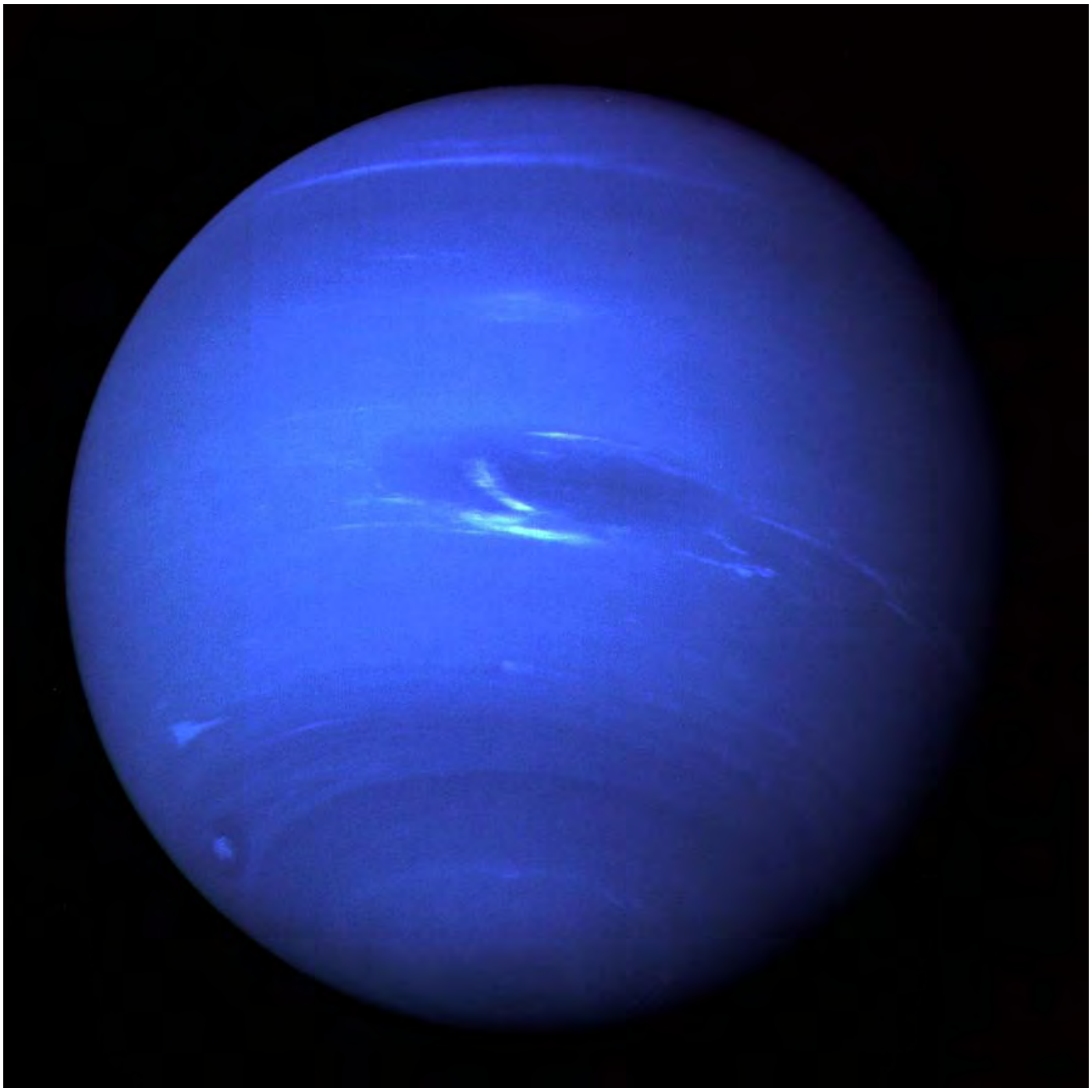


Figure 3.12: Neptune

3.4 Pluto and the Kuiper Belt

The ninth planet from the Sun, **Pluto**, is the odd-world out among the planets of the outer Solar System. It is far smaller than not just the gas giants, but also the rest of the planets including Mercury. With a diameter of 2300 km, it is even slightly smaller than our Moon. Its orbit is also highly inclined with respect to the rest of the planets; it is also more flattened and less circular—another way to describe it is that it has a high **eccentricity**—than the other planets. At its closest point in its orbit, it is slightly closer to the Sun than Neptune. Pluto also has a single satellite, **Charon**, which is about half the size of its parent planet. Charon is larger, relative to Pluto, than any other moon-planet—for instance, the Moon is only a quarter the size of the Earth. As a result, the Pluto-Charon system has often been referred to as a double planet.

Figure 3.13: Pluto and Charon

A Hubble Space Telescope image of Pluto and its moon Charon taken in February 21, 1994. This picture allowed astronomers to directly measure the sizes of the two objects, giving Pluto a diameter of 2320 kilometers and Charon a diameter of 1270 kilometers (or more than half that of its parent).

Based on its size and distance from the Sun, Pluto and its moon Charon, are thought to be similar in composition to the other small bodies orbiting the gas giants. They thus are icy objects, covered with impact craters. Observations with the Hubble Space Telescope show evidence of methane, nitrogen, ethane, and carbon monoxide ices on the surface.

Pluto may in fact be the largest object that has its origins in the **Kuiper Belt**, a large group of planetesimals that had been postulated in the 1950s, but the discoveries of which did not occur until the last decade. Astronomers had suspected that a vast population of objects lay beyond Neptune's orbit. Jupiter contains 318 Earth masses worth of material accreted from the solar nebula; Saturn contains 95 Earth masses; Uranus has 14 Earth masses, and Neptune has 17 Earth masses. Based on this trend, one would expect the amount of gas in the original solar nebula to decrease with increasing distance from the Sun, but not drop to zero suddenly. However Pluto is only 0.2% of the Earth's mass, so if there is missing material from the solar nebula, it could presumably have condensed into icy bodies, too small to be seen because of their distance from the Sun.

It was only in 1992 that instrumentation became sensitive enough that astronomers were able to discover the first Kuiper Belt object 1992 QB_1. Over the last decade, hundreds more objects have been discovered with orbital distances from 40 AUs out to 50–60 AUs. Some investigators have estimated that there must be up to 70,000 Kuiper Belt objects that remain to be discovered. In recent years, a flurry of announcements have been made of objects with sizes very close to that of Pluto.

Table 3.1: Pluto/Charon and the Kuiper Belt Objects

Object	Diameter (km)
Pluto	2320
Charon	1270
Sedna	< 1500?
2004 DW	1500
Quaoar	1200
Ixion	1065
2002 AW197	890
Varuna	900

Over time, it would not be surprising if a Kuiper Belt object was discovered that turned out to be larger than Pluto. Even today, there is some debate over whether Pluto should continue to hold onto its status as a planet, or whether it should be demoted to a Kuiper Belt planetesimal. However ever since its discovery nearly 75 years ago (by Clyde Tombaugh in 1930), we have lived with the idea that our Solar System has nine planets, which is difficult to dismiss even among research astronomers. If Pluto were discovered today, it certainly would not be called a planet. But because old habits die hard, Pluto will probably retain its planet label for the foreseeable future.

3.5 Comets

One final category of objects found in the outer Solar System are the comets. Comets were first noticed by ancient peoples because they appear seemingly out of nowhere,

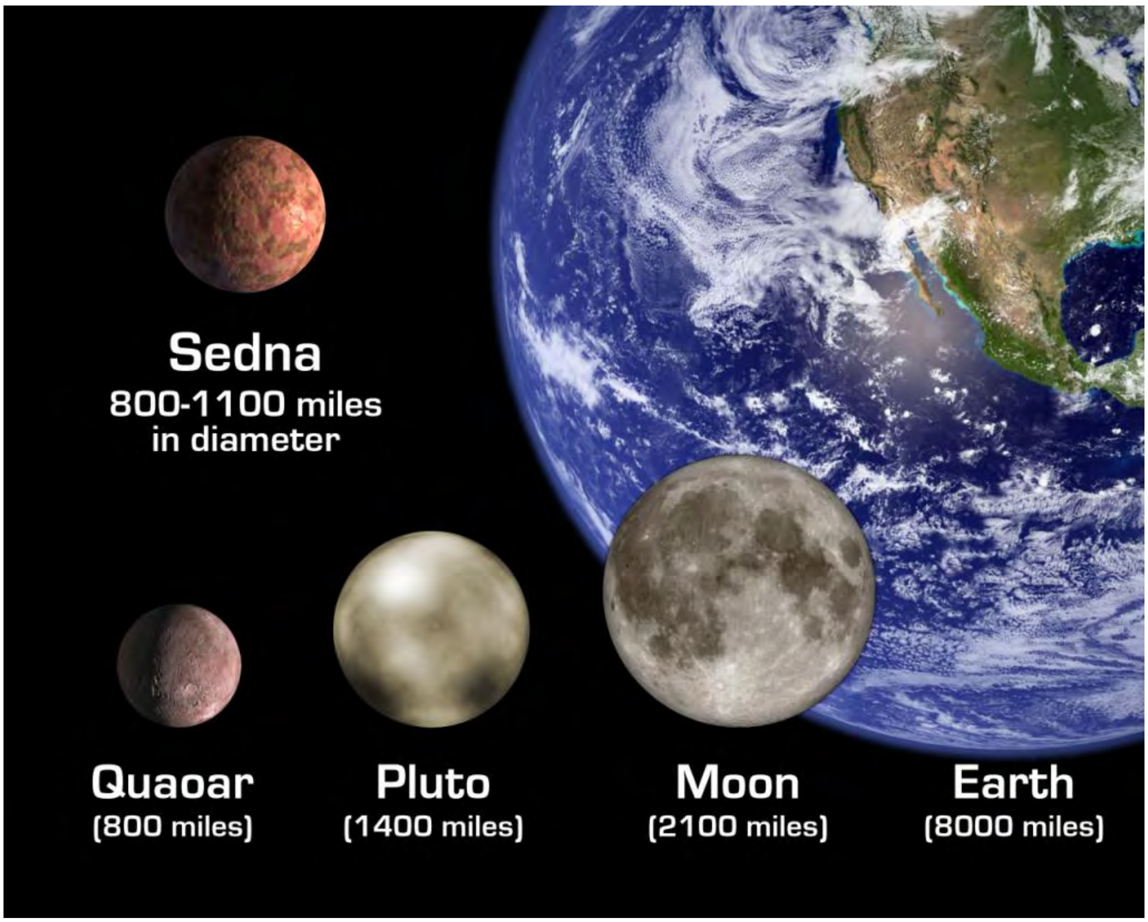


Figure 3.14: Comparison of Outer Solar System Bodies

Comparing the sizes of some recently discovered Outer Solar System icy objects to the Earth and the Moon. The large relative sizes of Sedna and Quaoar to Pluto has added fuel to the fiery debate over whether Pluto should keep its planetary status.

and drift between the stars, gradually growing larger and brighter. They do not "streak" across the sky, like meteors, but their motions are apparent after only a few hours or from night to night. Because there was no recurring pattern to their motions or appearance, comets were viewed by some as omens of impending doom and disaster.

Comets have several observable components: the bright diffuse center is the **comet head**, or **coma**. The **comet tail** can sometimes be faintly seen extending away from the head. Although comet tails can be seen extending a few degrees with the naked eye or through binoculars, long-exposure photographs can show them having lengths of tens of degrees. A telescope will also reveal a very bright, point-source center to the comet head, the **nucleus**.

We know today that comets are aptly described by the term "dirty snowballs." The nucleus of a comet is a sooty, icy world, that is anywhere from a few to several tens of kilometers across. Comets have highly elliptical orbits. According to Kepler's laws, they spend most of their time in the distant outer Solar System, and only a fraction of their time close to the Sun in their orbits. As they plunge into the inner Solar System, the energy from the Sun heats a comet's surface, *sublimating* (turning directly into gas) the icy volatiles. This gas along with the freed dust and other particulate material flows out from the nucleus and is driven back into the comet tail by the radiation pressure and streaming particles (the **solar wind**) from the Sun. A comet's tail therefore exists only when the comet is close to the Sun to be heated by it, and the tail will always point *away* from the Sun.

Because of their elliptical orbits, comets can be gravitationally nudged by planets that they pass by, especially the giant gas worlds. Over time as a comet makes repeated passages into the inner Solar System, its orbit may be modified by close encounters with the planets, until its orbit because less eccentric, more circular, and gradually smaller. Being much closer to the Sun, the comet's volatile ices will be lost as evaporating gas. A comet may dissipate completely until only the rocky solids and dust particles remain. The many meteor showers that the Earth passes through on an annual basis may be the dusty remnants of ancient comets that have long dispersed. Other comets might be redirected in their orbits so that they plunge into the Sun, or impact a planet. The SOHO spacecraft has observed dozens of comets destroyed each year as they fall into the Sun. A recent famous example of a comet hitting a planet is that of Shoemaker-Levy 9, which broke into nearly two dozen pieces before falling into Jupiter's atmosphere in 1994. On the Earth, the impact thought to have killed off the dinosaurs 65 million years ago may have been caused by a comet.

Based on the orbital characteristics of observed comets, there appear to be two distinct populations. **Short period** comets have orbits that take them perhaps a few hundred years or less to go around the Sun. **Long period** comets are far larger orbital radii, extending out thousands of AUs. The short period comets are thought to originate in the Kuiper Belt, where collisions between Kuiper Belt objects dislodge icy debris that can eventually rain down into the inner Solar System as short period

Figure 3.15: Comet C/2001 Q4

A comet discovered by NEAT, the Near Earth Asteroid Tracking program. At its bright center is the nucleus *of the comet. Surrounding it in a nebulous halo is the* coma. *The comet* tail *can extend for hundreds of thousands of kilometers.*

Figure 3.16: Nucleus of Halley's Comet
A close-up image of the nucleus of Halley's Comet as seen by the European spacecraft Giotto. Giotto made a closest approach of within 540 kilometers of the nucleus in March 13, 1986, and suffered heavy damage by high speed dust and ice particles in the coma.

comets. The long period comets are thought to originate in a completely separate reservoir, the **Oort Cloud**. This is a vast spheroidal region in the outermost boundaries of the Solar System, stretching from 50,000–150,000 AUs away, and containing as many as 100 billion comet nuclei. The comets in this location are so distant that they take a leisurely 10–60 million years to complete a single orbit about the Sun.

Although the Oort Cloud is 100 times further from us than the Kuiper Belt, current theories on the dynamics of the early Solar System suggest that the comets there actually formed much closer. Computer simulations show that icy objects forming in the first 100 million years after the collapse of the solar nebula would have been flung into the outer reaches of the Solar System by close gravitational encounters with the developing giant planets. Many would have been ejected completely out of the Solar System, but those that did not remained to form the Oort Cloud. A small fraction are gravitationally perturbed by passing stars each year. Some will get disturbed out of the Sun's gravitational influence entirely and be lost to interstellar space; others will be pushed inward, beginning a journey that could last thousands of years before they appear bright in our night sky.

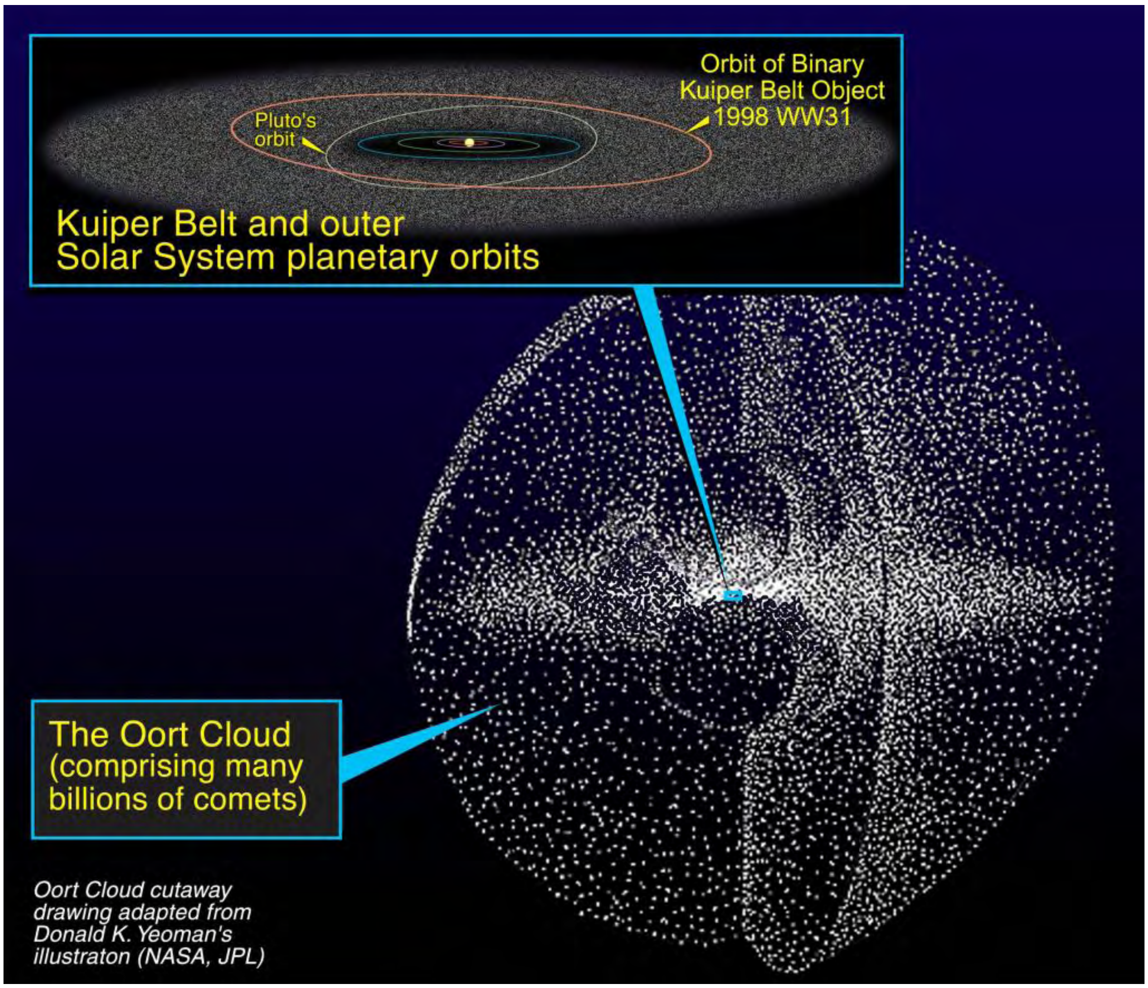


Figure 3.17: The Kuiper Belt and Oort Cloud

A comparison of the sizes of the Kuiper Belt, which extends out to roughly 60–70 AUs, and the hypothetical Oort Cloud, which is at least 100 times further away.

Lecture 4

The Sun

Stars, including our own Sun, are composed of gases. However they are so massive that their gravity compresses the gas atoms close together. Their internal temperatures are also great enough that the atoms are colliding constantly, with enough energy that at least one or more electrons are stripped from the atoms. Stars are 90% hydrogen and slightly less than 10% helium by number[1]. These two atoms have respectively one and two electrons each, which means the stellar interior is filled by an ionized gas, a dense soup of naked atomic nuclei swimming in a sea of free electrons[2]. This tightly packed ensemble makes the stars emit radiation in a manner similar to that of solids, that is, that of a blackbody curve. It is this fact which results in the primary method by which astronomers use to measure the temperatures of stars. They spread the star light with a spectrograph into a spectrum where the intensity at different wavelengths can be measured. These are compared to the theoretical blackbody curves of different temperatures. The curve that matches the measurements the best will have a temperature the same as the star being studied.

Stars however are complicated by having a thin gaseous atmosphere above the region where gravity is strong enough to compress the gas into a dense state. The gas in this region is far enough from the surface of the star that the temperatures are lower, so an atom might be ionized by having lost one electron, or not be ionized at all. These free atoms will emit and absorb light in the form of spectral lines. Light that escapes from the surface of a star passes through successively cooler layers of the stellar atmosphere. As a result, the spectrum that is measured on Earth from a star will consist of a continuous spectrum that originates from the dense, blackbody-emitting portion; the coolest part of the stellar atmosphere will have absorbed out specific spectral lines of the atoms found in this outer region. The stellar spectrum will therefore have dark **absorption lines** located where the emission lines for those atoms would normally be. These absorption lines were seen in the spectrum of the Sun as well as a star by the German physicist Josef von Fraunhofer in the early 19th

[1] With the rest of the elements making up one percent or less.

[2] The state of matter where most or all of the atoms are ionized is a **plasma**.

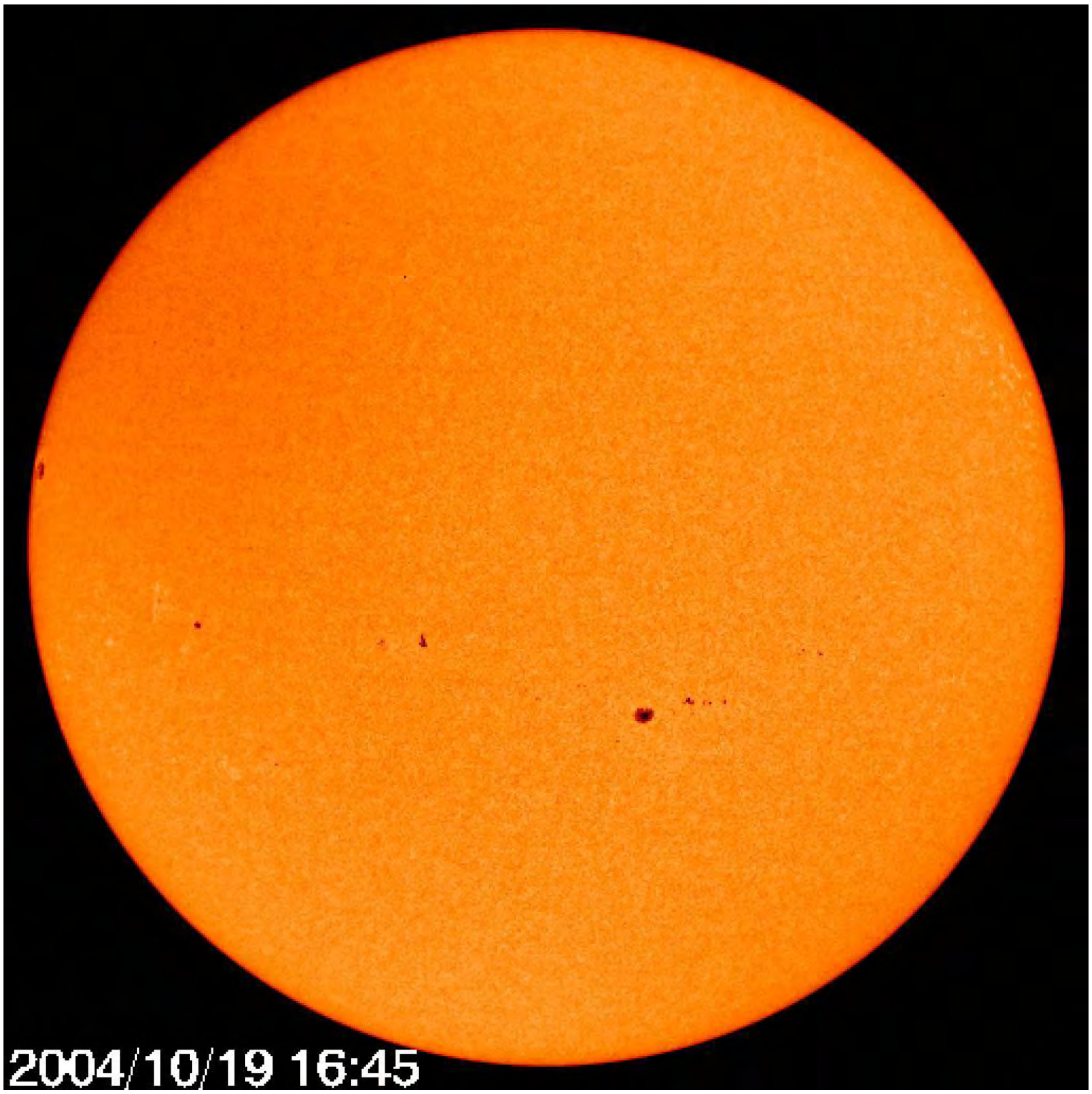
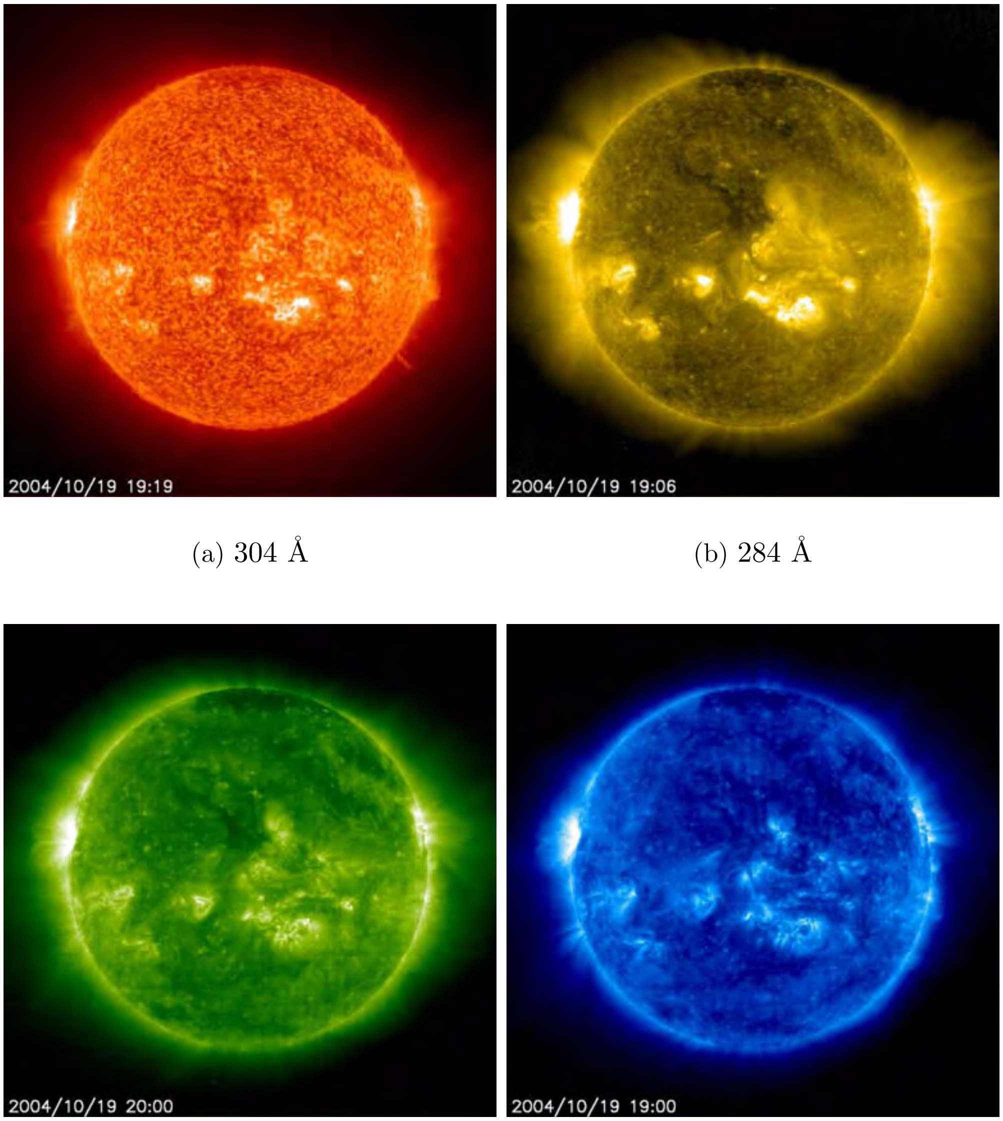


Figure 4.1: The Current Sun: Visible Wavelengths

A visible light image of the Sun taken the same day that this lecture was originally given in class. Just below right of center is a small clump of sunspots.



(a) 304 Å (b) 284 Å

(c) 195 Å (d) 171 Å

Figure 4.2: The Current Sun: X-Ray Wavelengths

The Sun as it appears in X-rays as seen by the SOHO spacecraft. These images were taken at roughly the same time as Fig. 4.1. The numbers refer to the wavelength of the X-ray emission, where the unit is an abbreviation for Ångstrom, a measure of length where $1\ \mathring{A} = 10^{-10}$ *meters. The different wavelengths trace different temperature gas, located from the photosphere to well into the corona. Note that the Sun appears much more active and violent in the X-rays than it does in the visible. Also the sunspot cluster visible in Figure 4.1 is bright in X-rays, and appears to have looping prominences of gas associated with it.*

century, and were subsequently named after him.

Table 4.1: Solar Composition

Element	Abundance (% of total number of atoms)	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	0.0030	0.14
Sulfur	0.0015	0.040

From these Fraunhofer lines, astronomers can determine the composition of the Sun. Tens of thousands of lines from 67 different elements have been detected. More elements probably exist within the Sun, but they are probably in too small of a quantity to show up with our present-day instrumentation. The ten most abundant elements on the Sun are shown in Table 4.1.

The compositions of the gas giant planets are very similar to the composition of the Sun, which is the leading piece of evidence suggesting that the Sun and its planets all form from the same original solar nebula. The compositions of other stars are also very similar, with large differences only found in the elements *other* than hydrogen and helium. This suggests a common origin for the gas that makes up all of the stars in our and other galaxies. We will see in §6.2.2 where a common origin for this gas leads us.

The structure of the Sun can be divided into roughly two regions, an interior that is effectively invisible and an exterior that we can observe. The density of this gas at the interior is so great that a photon cannot escape easily, but is scattered by particles in the solar plasma. The Sun's interior is therefore said to be *opaque*: the radiation cannot stream freely outward but bounces around from one nucleus to another, and between one electron and the next. At some point far enough from the center of the Sun, the density drops enough that the radiation *can* travel large distances without colliding with a atom or part of an atom. The region where this occurs is called the **photosphere**, from the Greek words meaning "sphere of light," because this is the portion of the Sun that is actually visible to us. Thus although the Sun is a giant ball of gas with no well defined "surface" such as what you would find on a terrestrial planet, we do *perceive* a layer that looks like a surface. The Sun is 700,000 kilometers in radius, and the photosphere is only 500 kilometers deep, or about 0.1% of the radius, As a result, the Sun appears to have a sharp boundary.

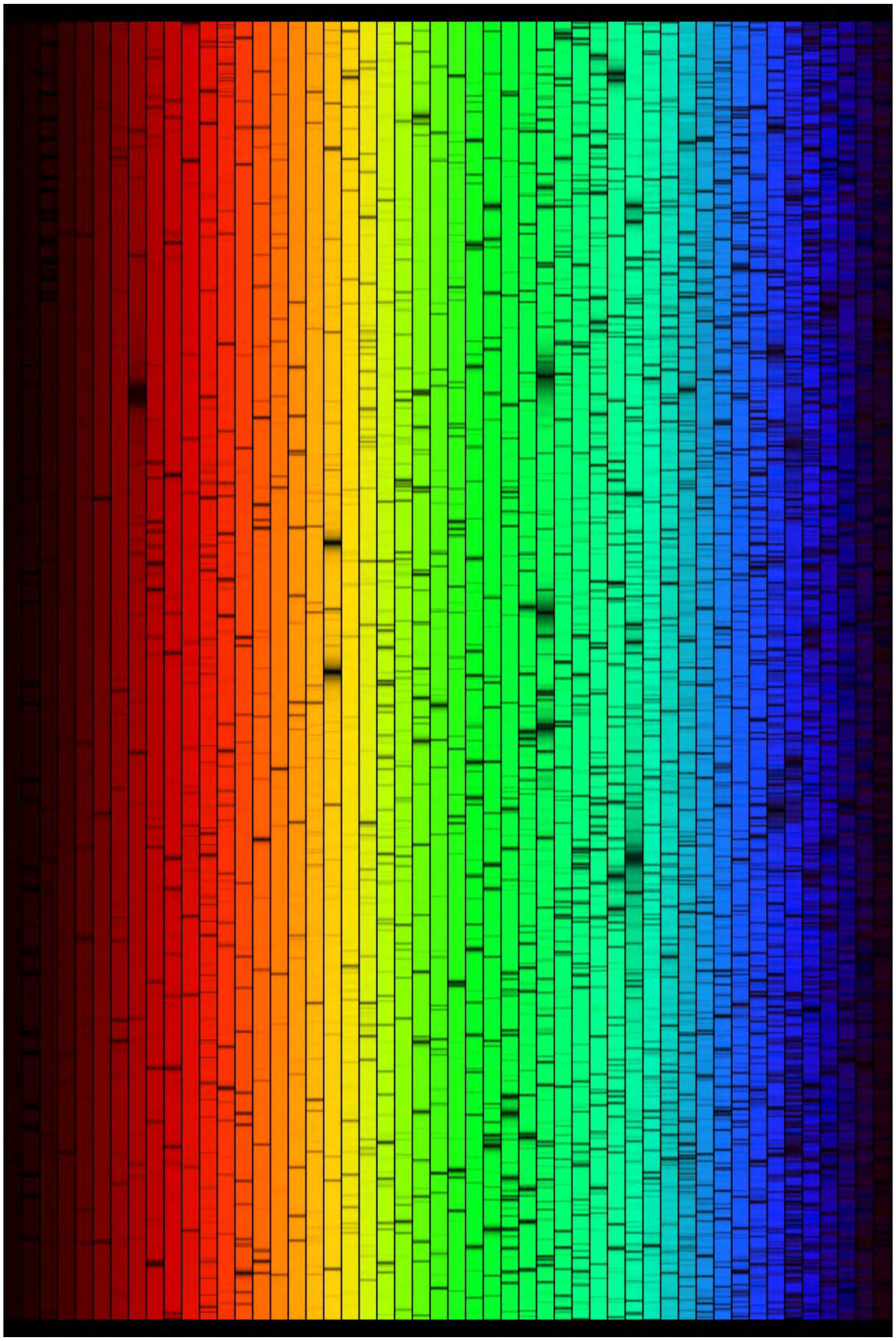


Figure 4.3: Fraunhofer Lines

A high resolution spectrograph of sunlight, spread out by wavelengths to show the locations of tens of thousands of individual absorption lines. These lines are formed by cooler atoms in the solar atmosphere that absorb the exact suite of wavelengths of light associated with each atom.

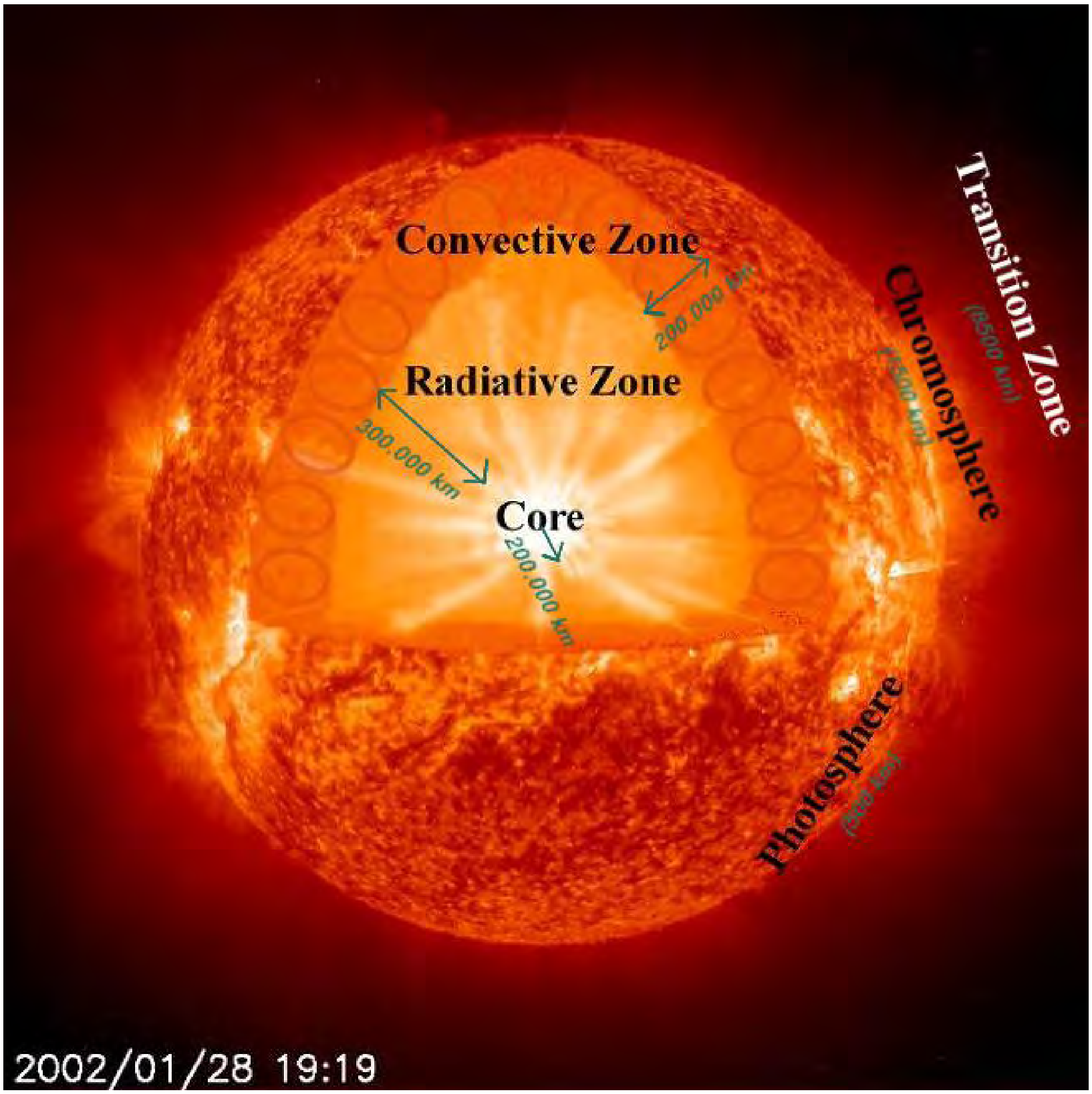


Figure 4.4: Structure of the Sun

All of the major interior and atmospheric components of the Sun are labeled except for the tenuous coronal gas, which is located just above the transition zone.

The interior of the Sun has several components. At the center is the solar **core** (about 200,000 kilometers in radius), where all of the energy is generated, and the temperature is at 15,000,000° K.[3] Next is the **radiation zone**, a layer 300,000 kilometers deep, where the temperature drops to 7,000,000° K. Above that is the **convection zone** only 200,000 kilometers deep, and extending out to the photosphere.

4.1 Nuclear Fusion at the Sun's Core

The energy from the solar interior is generated by **thermonuclear fusion**. The basic process is for two light nuclei to collide together to form a heavier third nucleus, and releasing energy in the process:

$$\text{Nucleus 1} + \text{Nucleus 2} \rightarrow \text{Nucleus 3} + \text{Energy}$$

The combined mass of Nucleus 1 and Nucleus 2 is actually *greater* than the mass of Nucleus 3. However Einstein showed from his theory of Special Relativity that mass and energy are equivalent, and one can be transformed into another. This is his famous $E = mc^2$ equation, which relates the energy E that is found in some matter of mass m, where c is the speed of light. The difference in mass between the starting and final nuclei in the reaction has been converted into energy.

Nuclear fusion in the core of the Sun is made possible only by the extreme temperature and pressure conditions in the solar interior. With sufficient energy, two individual protons (p) collide together to form a **deuteron** nucleus (d), a form of hydrogen with a neutron in addition to the proton:

$$\mathrm{p} + \mathrm{p} \rightarrow \mathrm{d} + e^+ + \nu,$$

a process which also releases a **positron** e^+ (the antimatter mirror to an electron e^-), and a particle called a **neutrino** ν. This is simply just the first step in the **proton-proton chain**, the fusion process in the Sun. The second step involves a deuteron and an additional proton, which collide to form Helium-3 (^{3}He), a form of helium with two protons and one neutron:

$$\mathrm{d} + \mathrm{p} \rightarrow {}^3\mathrm{He} + \text{Energy}$$

Finally the last step involves the collision of the two ^{3}He nuclei:

$$^3\mathrm{He} + {}^3\mathrm{He} \rightarrow {}^4\mathrm{He} + \text{Energy}.$$

The resulting Helium-4 contains two protons and two neutrons in its nucleus and is the type of helium most familiar to us. If we represent the single protons as ^{1}H (a

[3] The unit "K" stands for Kelvin degrees; see Appendix A for a further explanation.

normal hydrogen atom which has a single proton in its nucleus), then the net effect of these three separate steps can be written as:

$$^1\text{H} + {}^1\text{H} + {}^1\text{H} + {}^1\text{H} \rightarrow {}^4\text{He} + \nu + \nu + \text{Energy}.$$

The fundamental energy generation process at the core of the Sun is therefore *hydrogen being converted into helium, with the net release of energy.*

4.2 The Radiative and Convective Zones

The energy released at each step of the proton-proton chain is in the form of gamma rays. These photons slowly diffuse outward from the center of the Sun, absorbed and then re-emitted by the ions and electrons it encounters on the way. Through the core and the radiative zone (where the temperature is still high enough that virtually all the atoms are completely ionized), the radiation tends to move somewhat freely and only occasionally collides with a ionized nucleus or a free electron. Each collision results in a slight degradation of the energy of the photons as they pass through successively cooler layers on their way out of the Sun.

In the next zone outward from the center, the temperature has dropped to 2,000,000° K, which is sufficient for some of the larger atomic nuclei to start binding with free electrons. Instead of simply scattering off the nuclei and electrons, photons can get absorbed into the ionized atoms, causing a bound atom to jump into an excited state. The chance of photon capture is higher here, resulting in a *higher opacity* of the cooler plasma gas. The photons do not stream as easily as they do in the radiative zone below. As a result, the energy in a packet of gas builds up, raising its temperature, and causing a drop in its density. The gas packet rises, taking its built up energy with it, where it can be released higher up in the convective zone. The physical conditions in the Sun in this region are ripe for *convection*. This is the same process first discussed in §2.4.2, on the topic of heat transfer from the center of the Earth to its surface.

The convective currents in the Sun are quite complicated, with different tiers of convective cells, each with their own individual set of boiling motions, and each at a different depth in the Sun's interior. The largest cells are deepest, about 200,000 kilometers below the surface. These cells drive cells just above them which are not as large; those in turn drive medium-sized cells further up, and those drive smaller cells, continuing until at the very top of the photosphere, where the cells are about 1000 kilometers across and 1000 kilometers deep. These cells can be imaged by satellites and instruments on Earth as **granules** on the solar surface. The bright centers of the granules are regions where hot material is upwelling from below. This material emits its heat energy in the form of light into space, cools, and then sinks back down along the dark edges of the granules.

By the time the energy from the core arrives at the photosphere, the wavelength of the radiation has been downgraded until it is mostly in the optical. Because the Sun

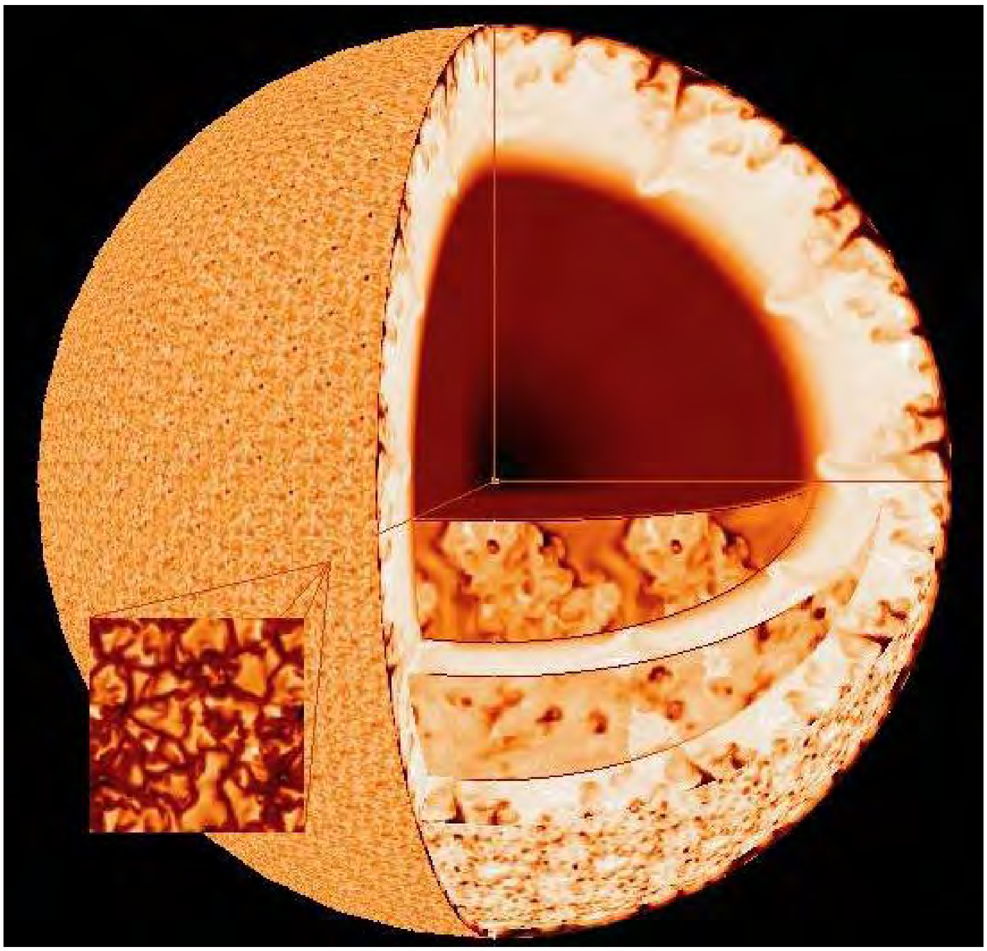


Figure 4.5: Convection Cells Inside the Sun

A computer simulation of the interior of the Sun's convection zone. The convection cells are largest in the deepest convection layer, and they grow subsequently smaller as we ascend up through to the surface. At the exposed photosphere are the smallest convection cells, the visible solar granules.

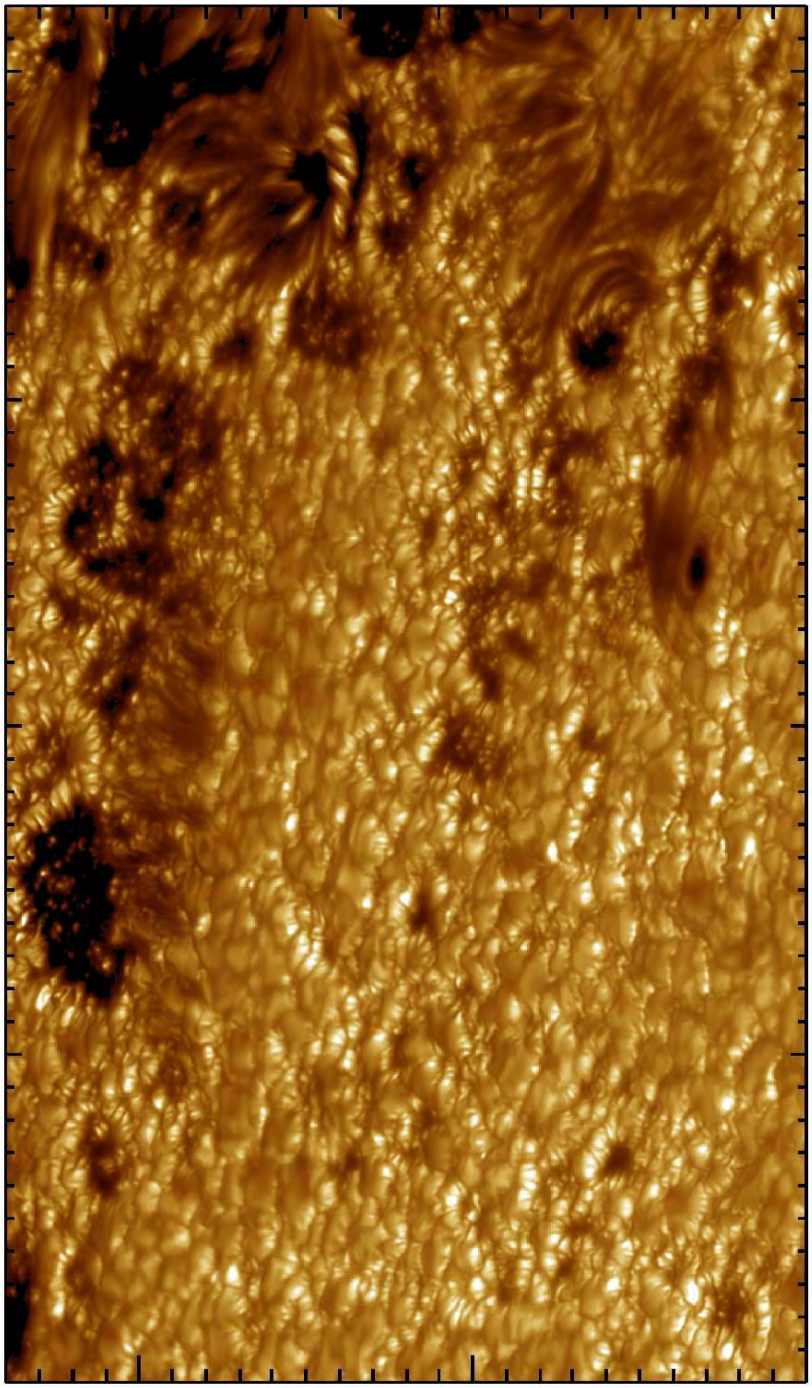


Figure 4.6: Granules in the Photosphere

An oblique view of the Sun taken from the Swedish Solar Telescope. The granular appearance is due to the churning of the topmost layer of convection cells just below the visible surface of the photosphere.

consists of a very dense gas, it (like all stars) emit as blackbodies. At the photosphere, the **effective temperature** of the Sun is equivalent to that of a blackbody at 5800° K.

The mass that is converted into energy in the proton-proton chain is small—the difference between the mass of the four starting protons and the final helium-4 nucleus is only 4.8×10^{-27} kilograms. However this means 1 kilogram of hydrogen will release 6.4×10^{14} Joules of energy according to Einstein's formula. A Joule (J) is a measure of energy; a more familiar use of this term is the Watt (W), which is the rate of energy produced over time; 1 W = 1 J per second. What this means is that 1 kg of hydrogen undergoing complete fusion will result in enough energy to run a single 100-Watt bulb for 20,000 years. Equivalently, the amount of energy generated inside the Sun *each second* is the same as that from 100 billion 1 megaton nuclear bombs going off. For the Sun to produce its current output of 3.85×10^{26} W of power, it consumes 600 million tons of hydrogen at its core every second. However the Sun is so massive that it has been burning its hydrogen for the past 5 billion years, and is thought to have enough fuel to continue its nuclear processes for another 5 billion years.

4.3 The Solar Atmosphere

Above the solar photosphere is the lower atmosphere, called the **chromosphere**. The density of the atoms in the chromosphere is too low for the chromosphere to be directly seen, especially if the photosphere is shining right below it. However the chromosphere can be detected in total solar eclipses, when the photospheric light is blocked. It glows with a red hue (hence the name "chromosphere" from the Greek words meaning "sphere of color"), due to the emission from the hydrogen alpha (Hα) emission line of atomic hydrogen. The chromosphere extends about 1000 km above the photosphere, where the temperature drops to 4500° K.

Above the chromosphere is a **transition zone** where the gas density continues to drop, but the temperature actually *rises*. It continues to rise until about 10,000 kilometers above the photosphere, where the temperature maxes out at 1,000,000° K. This region of the Sun's atmosphere is the **corona**, and it extends out another 10,000 kilometers or so. There is still much debate by astronomers as to why the corona is so hot. It must be receiving energy from the Sun somehow; current theories involve magnetic disturbances on the surface which transmit energy past the chromosphere and into the outer corona.

In addition to electromagnetic radiation, free protons and electrons constantly escape from the solar corona out into interplanetary space. This stream of particles makes up the **solar wind**. The light from the Sun travels (by definition) at the speed of light, so it arrives at the Earth 8 minutes after leaving the photosphere. The solar wind particles move slower, but at a still substantial velocity of about 500 kilometers per second. These particles take several days to reach the Earth. The Sun can go

Figure 4.7: The Solar Chromosphere
The chromosphere becomes visible at the edge of the Sun during a full solar eclipse. Normally the much brighter photosphere makes it impossible to see the chromosphere directly. The red emission comes from the de-excitation of hydrogen atoms.

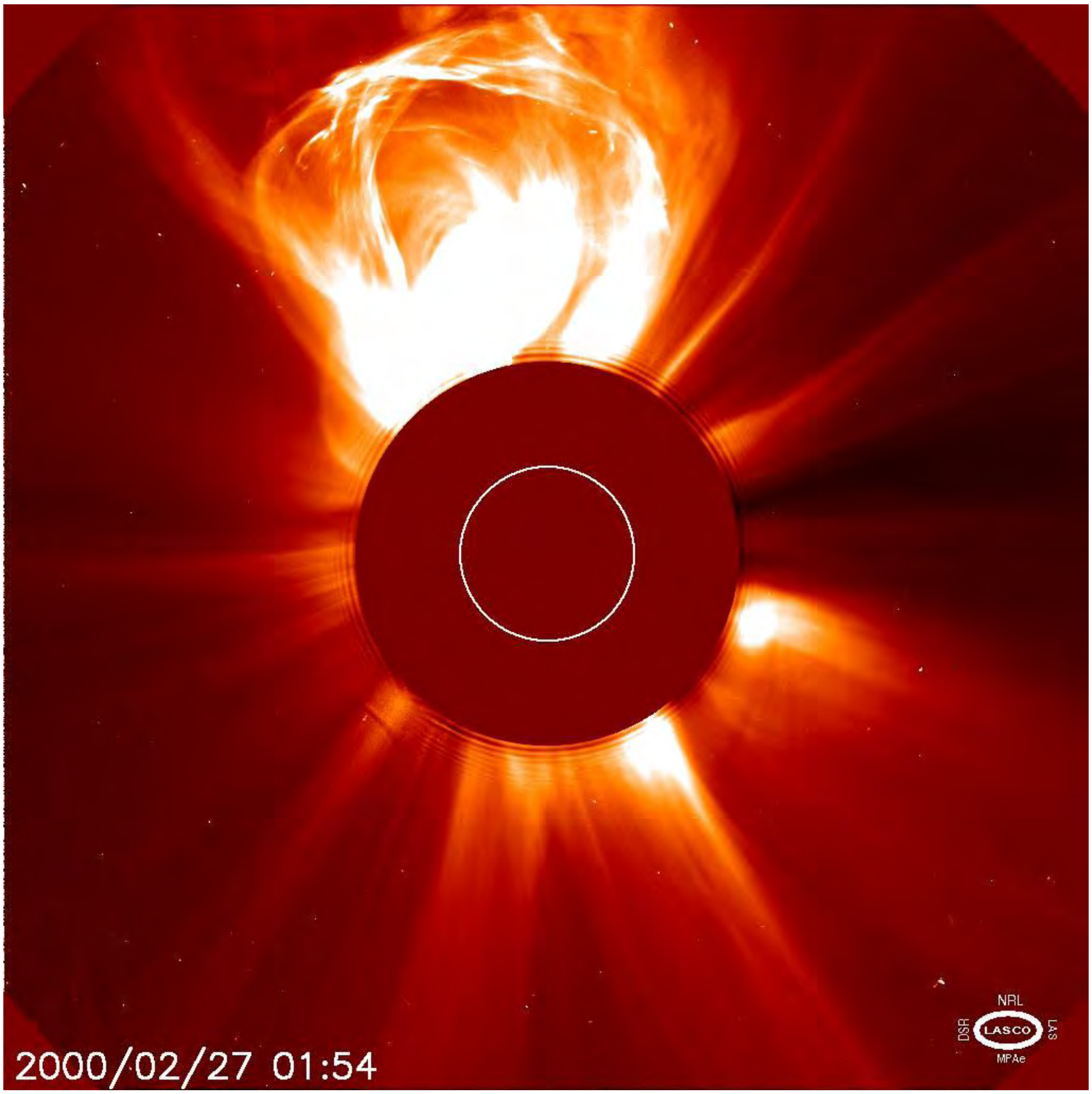


Figure 4.8: Coronal Mass Ejection

The corona becomes visible via an instrument called a coronagraph *on the SOHO spacecraft, which simulates solar eclipse by blocking the disk of the Sun. The coronal loops extend hundreds of thousands of kilometers into space and eject particles that eventually become part of the solar wind.*

through **active** as well as quiescent phases. When it is active, the surface of the Sun is much more violent, with more dark **sunspots** appearing prominently on the solar disk; solar **prominences**, loops and sheets of hot gas erupting out of the surface; and solar **flares**, even more violent releases of energy from small pockets of gas with temperatures up to 100,000,000° K. All of this activity pumps up the corona as well, with the energy eventually finding its way to more energetic as well as a greater number of particles in the solar wind. The particles that result from flare activity can directly impact us when they arrive at the Earth. They light up the atmosphere near the north (and south) poles to create the aurora borealis (and aurora australis); affect power grids on the surface; and disrupt communications satellites in orbit above it.

Lecture 5

Stars and the Milky Way Galaxy

5.1 The Life Cycle of Stars

Our Sun resides in the Milky Way galaxy, a collection of about 200–300 billion stars, bound together by each others' collective gravitational pull. In Lecture 2, we had discussed the origins of the Solar System from a pre-solar nebula. But where did this nebula originate from? To answer this question, we will cover in this Lecture the stars, the basic building block of galaxies, and understand their origins and their life cycle, and along the way, review many of the other astronomical phenomena that can be found within a galaxy.

Because we live inside the Milky Way, it is difficult to imagine what the actual shape and size of our parent galaxy is. Today we know that it belongs to a class of objects known as **spiral galaxies** (and more specifically, a *barred* spiral galaxy). We will discuss the variety of other types of galaxies in the next and final Lecture. A galaxy that might look similar to ours is M 51. A spiral galaxy is distinguished by a large flat, **disk** of stars, with a central **bulge** at the center[1]. The characteristics of the bulge and the disk are very different: the bulge tends to have stars that are orange-ish in color, and is more uniform in texture. The disk, on the other hand, has **spiral arms** that run through it pinwheel-like. The arms themselves are clumpy, with brighter regions, as well as finer dark lanes. With filters that are sensitive to the Hα line of hydrogen (the same easily visible emission found in the Sun's chromosphere; §4.3), bright red knots of emission are also found along the spiral arms. Because of the predominance of the form of ionized hydrogen, these knots are known as H II regions.

The features visible in other galaxies turn out to have to do with the different phases in the life cycle of stars. The dark lanes along the spiral arms are enormous clouds of **interstellar dust** and gas. The gas collected in these clouds is cold enough that it is mostly molecular in nature; the hydrogen is not in atomic form (H), but

[1]For a culinary analogy, think of a spiral galaxy as a pizza with two hamburger buns stuck in the middle, one above and the other below the pie.

Figure 5.1: M51
*The central region of the spiral galaxy M51 as seen by the Hubble Space Telescope. This is similar to what the Milky Way would like like from the outside. This view also shows many of the steps in the star formation process, including the vast dark dust lanes and molecular clouds where the raw materials of stars are found, the H*II *regions where OB stars are tearing apart the molecular cloud where they are born, and the bright young clusters of stars which have not wandered far from the natal environments.*

is bound together two to an atom, in molecular form (H_2). These dark dusty clouds are therefore known as **molecular clouds**. The ones that are visible in M 51 are enormous: at least hundreds of light years across. They are **giant molecular clouds** or GMCs for short.

It is inside such molecular clouds where stars are born. However as we will soon learn, stars tend not to form just singly or even a few at a time. They form in clusters, with memberships numbering from the hundreds for the medium-sized clouds to tens of thousands for the largest molecular clouds. Once they form, the largest stars, which pump out copious amounts of ultraviolet radiation, light up the natal environment, creating the H II regions. As we will see, the radiation and the energetic winds from these stars disrupt and destroy the parent molecular cloud. The stellar cluster is freed from the gravitational confines of the molecular gas. As the individual stars orbit the galaxy, they gradually disperse, spreading out to fill the galactic disk.

The most massive stars do not have much of a chance to wander very far. They live at most 10 million years before dying in cataclysmic **supernova** explosions. These release enormous amounts of energy—about 10^{43} Joules all at once. These explosions can further destabilize and disrupt the nearby molecular clouds, pushing into them and plowing the gas outward to create giant **shells** and **superbubbles**, which are perhaps visible to observers from outside the galaxy. The violent ends of the massive stars along with the more peaceful deaths of less massive stars leave burnt out stellar cinders, compact objects with bizarre properties—the **white dwarfs**, **neutron stars**, and **black holes**. However the explosive supernovae also inject gas back into the **interstellar medium** in the disk of spiral galaxies. This gas is enriched by heavy elements, newly created inside the massive star during its life as well as during the supernova explosion. This gas eventually finds its way back into the giant molecular clouds, where the star life cycle begin anew.

5.2 Stars: Building Blocks of Galaxies

When we attempt to perform a census of our galaxy, we find that the stars make up the bulk of the ordinary matter. There is roughly 500 billion solar masses = 5×10^{11} $M_\odot$ contained within all of the stars in our Milky Way[2]. The space between the stars is not empty; this **interstellar medium** contains at least a warm (several 100° K) atomic gas with a minimum density of a few particles per cubic centimeter. Other regions of space might contain denser, cold molecular clouds, or hot plasma gas released from supernovae explosions. Despite a volume that is far greater than the total volume of stars, the gas component of the Galaxy is very tenuous. All of the different types of gas add up to only 5 billion solar masses of material, or just 1% of the stars' mass. Stars can therefore be considered the fundamental building blocks of galaxies.

[2]One solar mass is the equivalent amount of mass as that contained in our Sun, or 1.99×10^{30} kilograms.

Not all stars are the same however. They come in a range of masses with varying surface temperatures, sizes, luminosities, and life histories. When astronomers first started studying stars using spectroscopy in the late 19th century, they were able to identify some of the spectral lines with gases in Earth-bound laboratories. They initially classified stars based on the intensity of the hydrogen absorption features, with the letters A, B, C, D, ... This was before a full understanding of atomic structure and atomic spectra was possible with the advent of quantum mechanics in the 1920s. Using quantum theory, astronomers realized that stars were better classified by their surface temperatures. However the old letter classifications remained and were retained. We still use these letters although their sequence is jumbled because of this historical accident.

The modern method of stellar classification puts stars into the **spectral classes** (or **spectral types**) **O**, **B**, **A**, **G**, **K**, and **M**, where O stars are the most massive the spectral classes are in order of decreasing mass. One way to remember this sequence is to use a mnemonic, such as the famous:

Oh **B**e **A** **F**ine **G**irl, **K**iss **M**e.

or the alternate:

Oh **B**e **A** **F**ine **G**uy, **K**iss **M**e.

The spectral classifications are further subdivided into finer categories. For instance, the G class is broken into G0, G1, G2, ..., G8, G9, with a G9 star being the least and G0 being the most massive. The next largest star after a G0 is the smallest star in the F spectral class, F9.

When stars are arranged in order of decreasing *mass*, such as the preceding OBAFGKM, they will also be arranged in decreasing *effective temperature*, *decreasing size*, *increasing life expectancies*, and *decreasing rareness*. This is summarized in Table 5.1. The most massive stars are those with the O spectral type, while the least massive stars are in the M spectral class. Our Sun is toward the low end of the mass range, placed in the G spectral class.

Table 5.1: Stellar Properties

Spectral Class	Mass [$M_\odot$]	Luminosity [$L_\odot$]	Surface Temperature [° K]	Life Expectancy	How Common?
O	> 24	3,000–100,000	$> 30,000$	< 2 million	1
B	3—24	75–3000	20,000	30 million	250
A	1.7—3	10–75	10,000	400 million	1,750
F	1.2—1.7	2,5–10	7,000	4 billion	4,900
G	0.8—1.2	0.3–2.2	6,000	9 billion	12,500

Continued on next page

Continued from previous page

Spectral Class	Mass [$M_\odot$]	Luminosity [$L_\odot$]	Surface Temperature [° K]	Life Expectancy	How Common?
K	0.6—0.8	0.035–0.3	$4,000$	60 billion	$27,500$
M	0.08—0.6	2×10^{-7}–0.026	$3,000$	200 billion	$250,000$

Our Sun by definition has a mass equivalent to one **solar mass**, or 1 $M_\odot$ (the "$\odot$" symbol represents the Sun). The least massive stars are the M stars, and smallest of these have the minimum mass necessary for sustained hydrogen nuclear fusion at their cores. The O stars are at least 24 $M_\odot$ in mass, but there is some uncertainty to their maximum mass; it may be as much as 100 $M_\odot$, and there is some observational evidence that such massive stars do exist. Note that the range of masses (and luminosities) shown for each line in Table 5.1 represent the differences of the subdivisions within each spectral class.

The luminosities in the table are also listed in reference to the **solar luminosity**, which is defined to be 1 $L_\odot$. Although the O stars are no more than 100 times the mass of the Sun, they can have luminosities as great as 100,000 times that of the Sun's. And conversely, even though an M star is no smaller than 3% the mass of the Sun, the lowest luminosity that such a star can have is less than a millionth of a solar luminosity. The dramatic increase in energy production in O stars (and the opposite for M stars) is due to the sensitivity of the core nuclear reactions to temperature. The more massive a star, the greater the mass of material pressing down onto its core. The weight of this material increases the pressure of the gas at the core, which results in a correspondingly greater core temperature. Even a slight increase in the temperature will result in a much greater energy output.

Thus although the center of Sirius (an A star) is thought to have a temperature of 20 million °K, and is only 5 million °K hotter than the center of the Sun, Sirius is 40× more luminous than the Sun (a G star). When the radiation diffuses out from the center of the Sun to its surface, the gamma rays have been downgraded enough that the temperature of the solar photosphere (its effective temperature; see §4.2) is 5800° K. For Sirius, the total number of photons is greater and their energies are higher on average, so that the surface temperature is 10,000° K.

Because the most massive stars in the O and B spectral types have much higher nuclear reaction rates, they use up the hydrogen fuel in their cores at faster rates as well. Massive stars therefore live much shorter lives, despite their greater mass. One star that is more massive than another will always as well have a shorter **main sequence** lifetime, the time during which normal hydrogen fusion takes place. For stars far less massive than our Sun, the exact opposite is true. The nuclear fusion rates can be thousands of times lower, resulting in estimated lifetimes of hundreds of billions of years for the tiniest M stars, much longer than the age of the current universe. We can therefore apply the adage that *the star that burns brighter, burns* *faster.*

One final important difference among stars with different masses are the chances of finding them in the Galaxy. The last column in Table 5.1 lists the relative numbers of how many of each type of star you would expect to encounter in the Galaxy. Thus for every 1 O-type star, there will be 12,500 G stars, and a whopping 250,000 M stars. *High mass stars are rare, whereas low mass stars are very common.*

5.3 Star Formation

Stars form from clouds of cold molecular gas. The largest of these are the **giant molecular clouds** or **GMCs**, which are some of the largest structures in the Galaxy. They have sizes ranging from tens to hundreds of light years across, and possess up to several million solar masses worth of gas. Once these clouds of gas collect together and concentrate, the gas deep within the cloud core is sheltered from stellar radiation and other sources of heating. These clouds therefore tend to be very cold, typically with temperatures less than 100° K, and for the densest, deepest cores, temperatures between 5°–10° K.

Because the temperatures are so low, the composition of GMCs is primarily molecular; instead of atomic hydrogen (H), molecular hydrogen (H_2) is the most common, making up 70% of the mass of the cloud. Helium makes up 29% by mass, while the remainder of 1% is taken up by **trace molecules** such as carbon monoxide (CO), carbon monosulfide (CS), hydroxyl (OH), and many other more complex molecules.

Despite their low concentrations in GMCs, some of these molecules, especially CO, are important for astronomers in studying cold clouds. Molecular hydrogen turns out to be extraordinarily difficult to detect, at GMC temperatures. Other molecular species like CO are far easier to excite and give off emission lines. It is primarily via such lines from CO that astronomers have learned so much about molecular clouds.

Radio telescopes can in fact map entire GMCs using the CO line, and what they have found are structures that are surprisingly complex. Instead of just a simple blob of cold, dark gas floating in space (although such simple clouds do exist), the majority of GMCs are sprawling structures with fine detail appearing at many scales. GMCs have structures that can be described as filaments, shells, sheets, columns, in addition to regular globules. There is fine detail at multiple scales: they appear to be fractal-like.

The densest parts, the **cores** of GMCs, also tend to be regions where young stars are found. This makes sense since it is only in such cold regions where the gas is dense enough that it might collapse into stars. The initial collapse probably does take some kind of trigger, such as a shockwave from a nearby supernova or powerful winds from other stars impacting the outer boundaries of the cloud. Once a cloud core starts collapsing, it does so **hierarchically**. The single cloud core breaks up into a handful of sub-cores which continue to collapse. These sub-cores can also break apart into even smaller cores, and so on, until the mass of the gas collapsing inward is

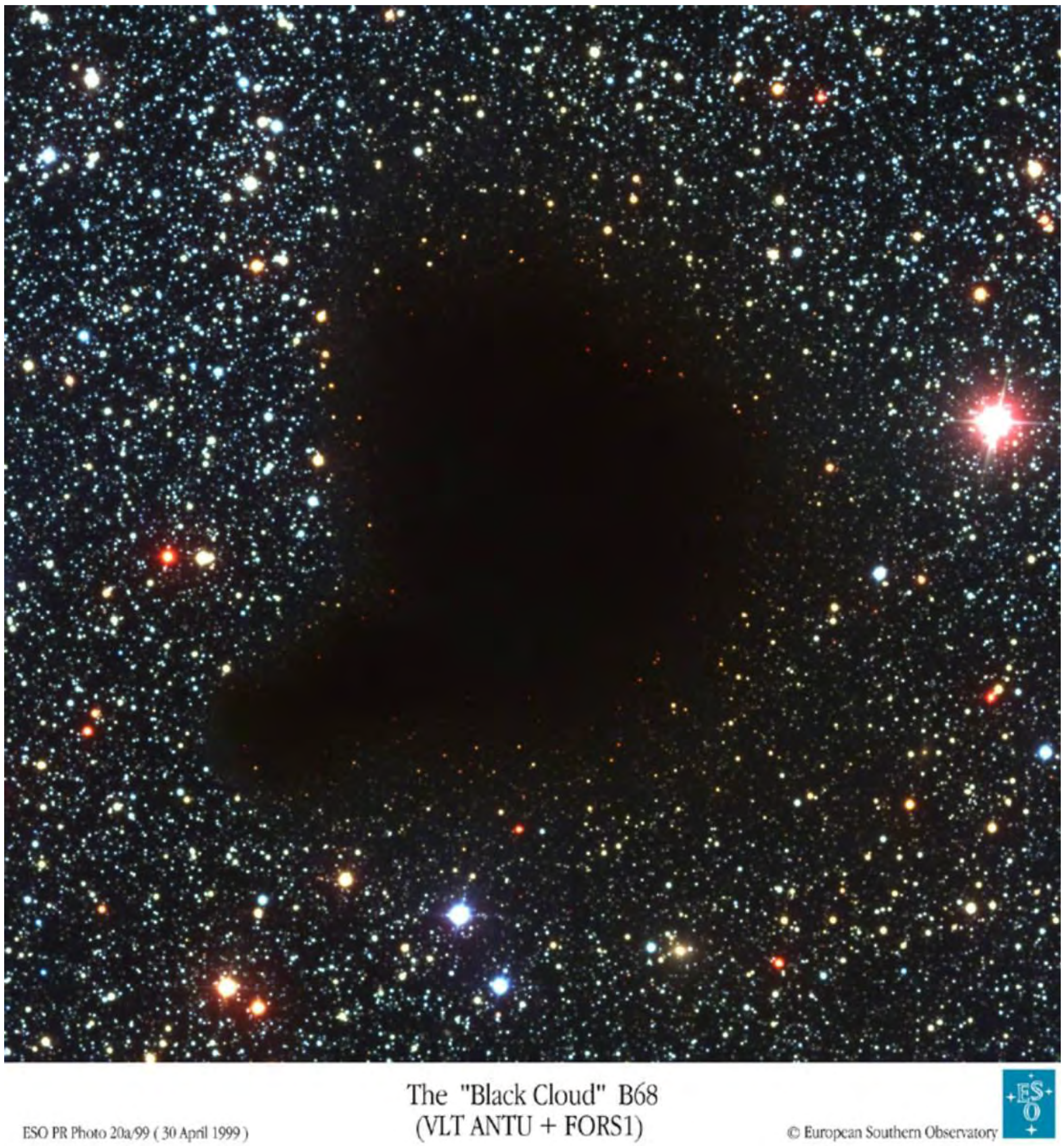


Figure 5.2: Barnard 68

An example of a solitary molecular cloud. The dust and gas in this cloud is so dense that it blocks all background starlight from reaching us. No stars are currently forming within this cloud.

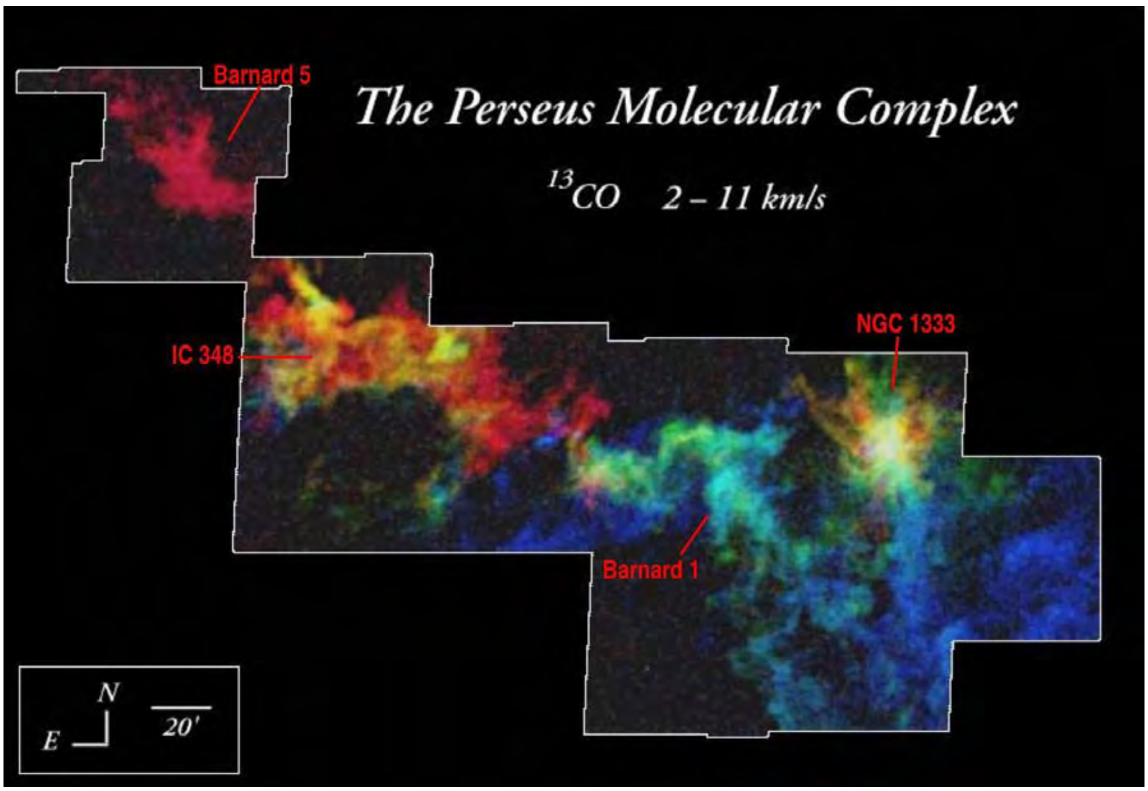


Figure 5.3: The Perseus Molecular Cloud Complex
A giant molecular cloud complex in the constellation of Perseus. This image was built up from tens of thousands of microwave radio observations observing an emission line from ^{13}CO. *This version of carbon monoxide has a carbon atom with an extra neutron in its nucleus, giving it 6 protons and 7 neutrons at its center (13 total) as opposed to the more common carbon-12 form.*

maybe a few times that of a single star. It is thought that the differences in masses between the sub-sub-cloud cores eventually leads to the different numbers of stars of different spectral types that we observe in the Galaxy. The scenario that also emerges from this picture is that stars are not formed in isolated environments. Because the original cloud core has undergone multiple splits, the final result is a cluster of stars as opposed to a solitary protostar.

It is one of these smallest sub-cores of collapsing gas that would have turned into the solar nebula described in §2.3. Much of the gas falling in ends up in the central star, but much of it is also *ejected back out into space* via a pair of jets that point in opposite directions along the north and south poles of the accreting **protostar**. This seems counter-intuitive—why should material falling in toward a star get expelled again? There is still much debate over the exact mechanisms by which this happens, but most involve interactions between the nascent magnetic field from the protostar and the infalling material which has gathered into an **accretion disk** around the center. What is clear however is that a protostar actually injects energy back into the molecular cloud via these jets of fast moving gas[3].

Less than 0.1% of all stars formed end up being massive O and B stars (which we will collectively call **OB stars**). Sufficiently large GMCs such as the clouds in the Sword of Orion form hundreds of thousands of stars, so that chances are, a handful of OB stars will form as well. The most massive of GMCs could form dozens of OB stars. Once these stars "turn on," their surface temperatures are high enough that the majority of their radiation output is in the ultraviolet. This radiation along with powerful winds emanating from the same stars slam into the surrounding molecular cloud gas, splitting each molecule of hydrogen into two hydrogen atoms and subsequently ionizing each atom by stripping the lone orbiting electron away from the hydrogen nucleus.

As soon as OB stars come into existence, they quickly begin to destroy their natal environment. The irradiated hydrogen gas forms a hot bubble that grows over time, at the expense of the cool molecular gas that makes up the rest of the GMC. When these bubbles of 10,000° K gas break out of the GMC, they reveal themselves to us as H II regions, such as the Orion Nebula in the Sword of Orion[4]. In images of the Orion Nebula and other H II regions, we are often staring into a cavity in the GMC that has been carved out by the radiation and winds from the OB stars.

Even GMCs that do not have any OB stars are slowly disrupted by the star formation process, since even the low mass stars will send out powerful jets that can rend apart the surrounding cloud. The Perseus molecular cloud complex has a number of small cores, such as NGC 1333 where dozens of low mass protostars are

[3]The speeds vary over time, but even the youngest protostars can expel jets at 10–20 kilometers per second; the jets from more mature protostars can travel at hundreds of kilometers per second.

[4]Before an H II region erupts to the surface of a GMC, they are invisible at visible light wavelengths. However they are bright in the radio and many of these **ultracompact H II regions** have been found from radio surveys.

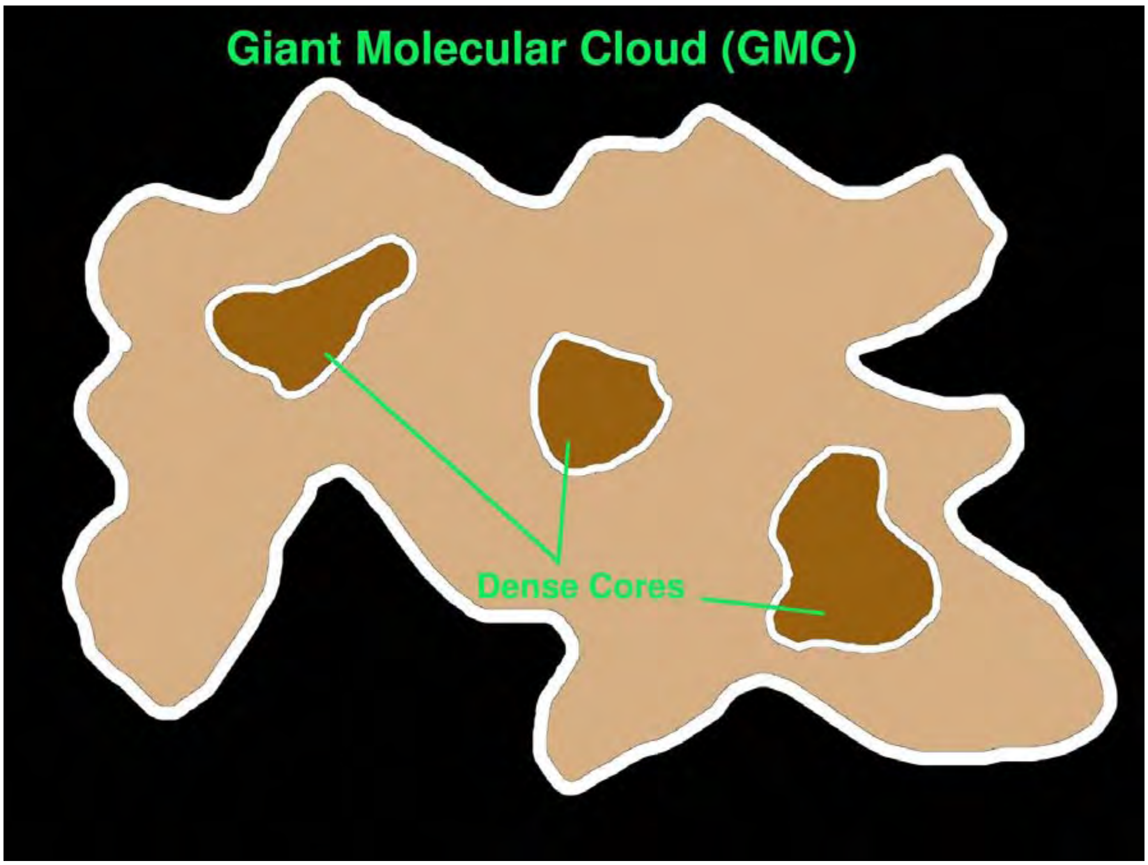


Figure 5.4: Star Formation: The Giant Molecular Cloud
A giant molecular cloud (GMC) about to form stars has dense cores that are critically near collapse. They only need a small external nudge, such as a supernova shockwave or winds from nearby stars, to initiate collapse.

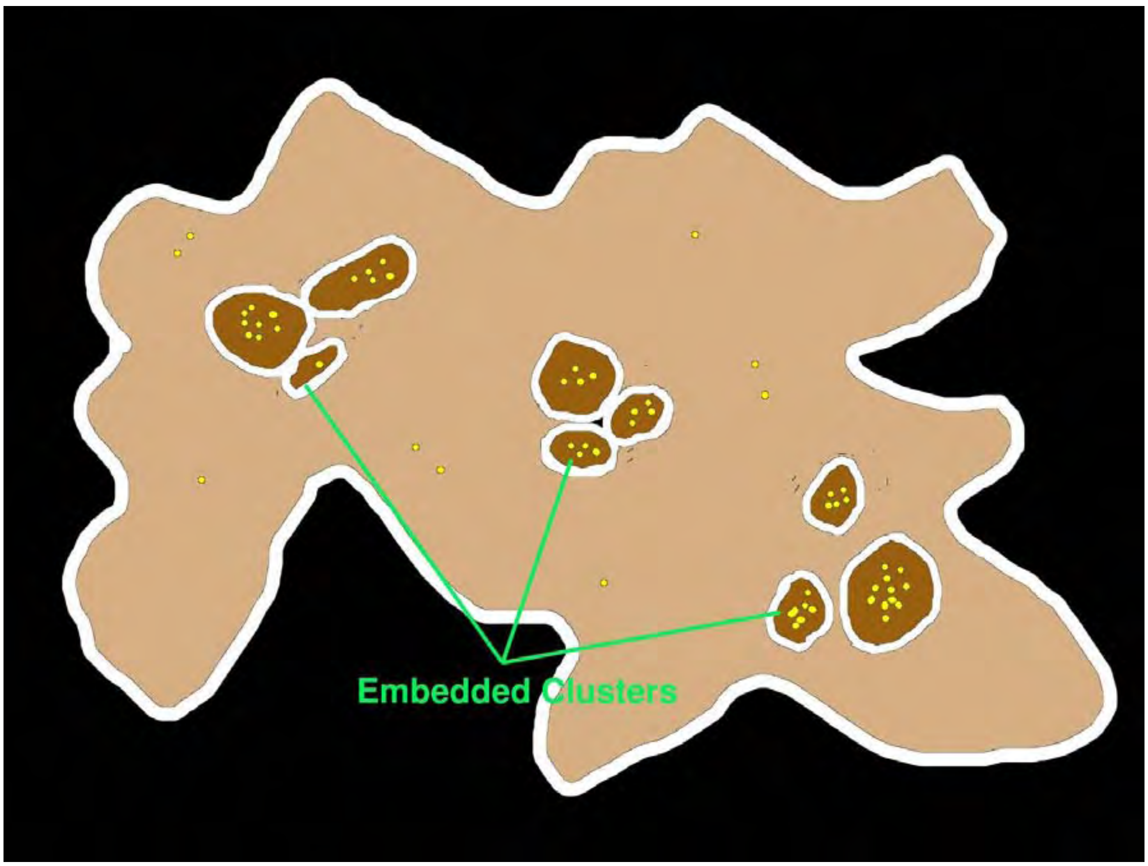


Figure 5.5: Star Formation: Embedded Clusters

After the cores start collapsing inward, they break up into smaller sub-cores and even smaller sub-sub-cores, until the cloud fragments are roughly the mass of the single stars. The stars formed in this manner tend to be highly clustered, so that dozens of newly born stars are found within the vicinity of the original dense cores.

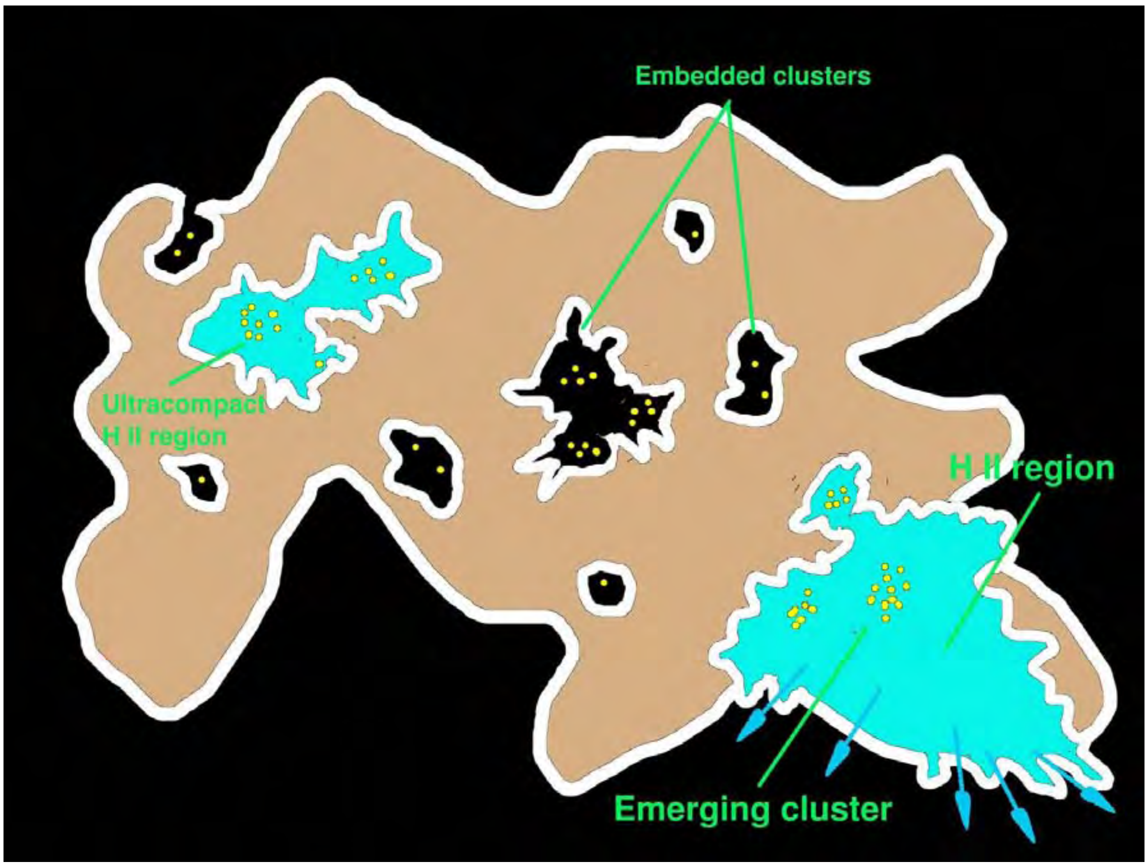


Figure 5.6: Star Formation: Emerging Clusters

Star formation is an inherently destructive phenomena. If OB stars are born inside a cluster, the UV radiation from these massive stars immediately eat away the surrounding molecular gas, and fill the expanding cavity with hot, plasma gas. Such embedded clusters inside hot bubbles of gas can be detected in the radio, and are known as ultracompact H II regions. *When this expanding bubble of gas breaks through the surface of the molecular cloud, the cluster of protostars becomes exposed and we have an H*II *region as a result. Even embedded clusters without massive OB stars slowly churn and destroy the molecular cloud from within.*

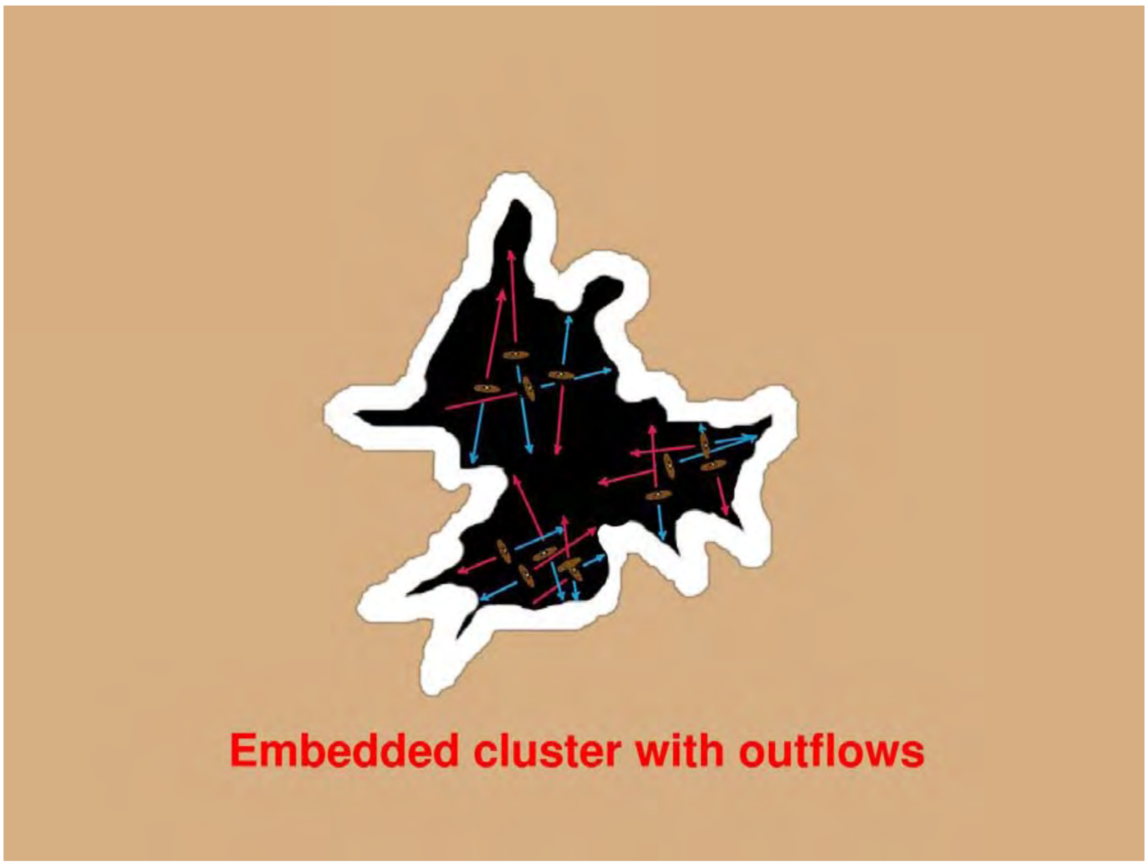


Figure 5.7: Star Formation: Low Mass Stars with Outflows
An embedded cluster without any massive stars will still destroy surrounding molecular gas. Even the lowest mass protostars have bipolar outflows which emerge from their north and south poles. Gas is ejected from these outflows at 100s to 1000s of kilometers per second.

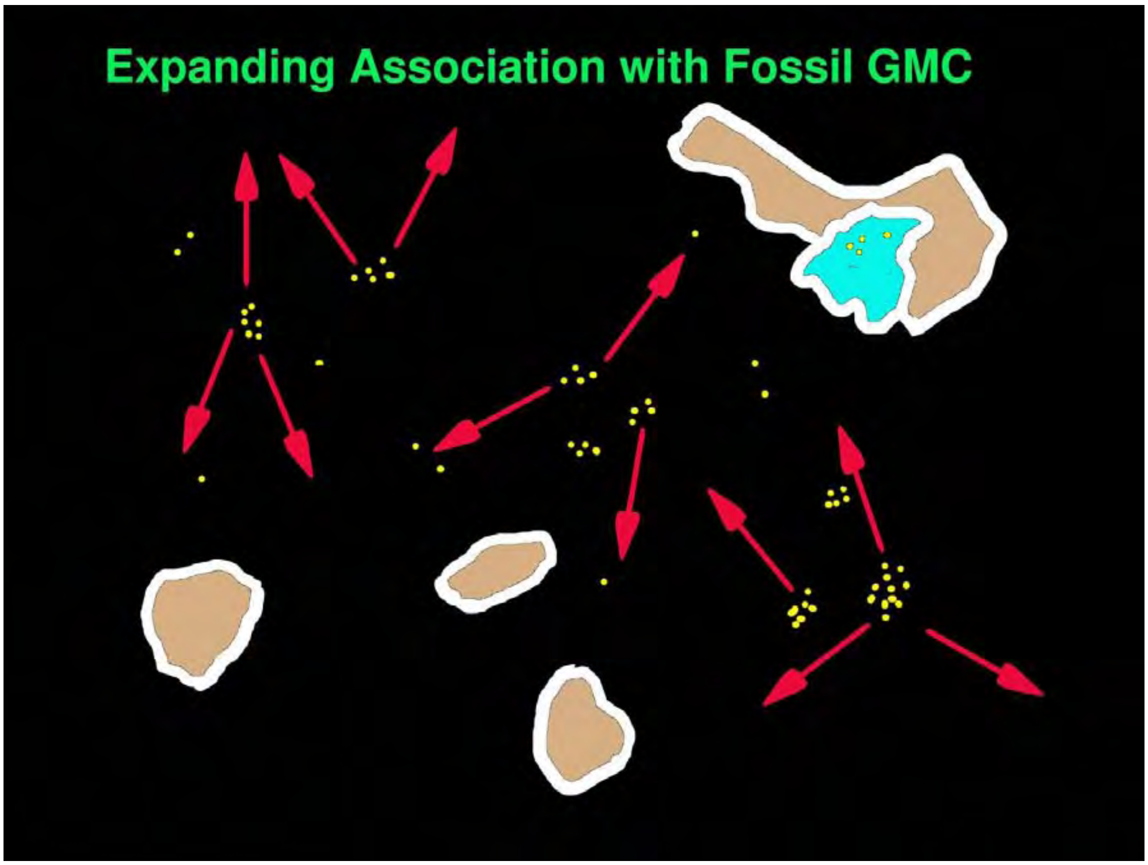


Figure 5.8: Star Formation: The End of the GMC

After several million years, the original giant molecular cloud has been mostly destroyed by the star formation that took place inside it. The young clusters of stars are no longer gravitationally bound by any surrounding gas so they slowly expand through the Milky Way as a loose "association." Star formation may also be self-propagating: in the upper right, the winds and outflows from protostars have caused one corner of the GMC to collapse and form a new cluster of protostars.

Figure 5.9: BHR 71
An outflow from an embedded young stellar object emerges from this isolated molecular cloud.

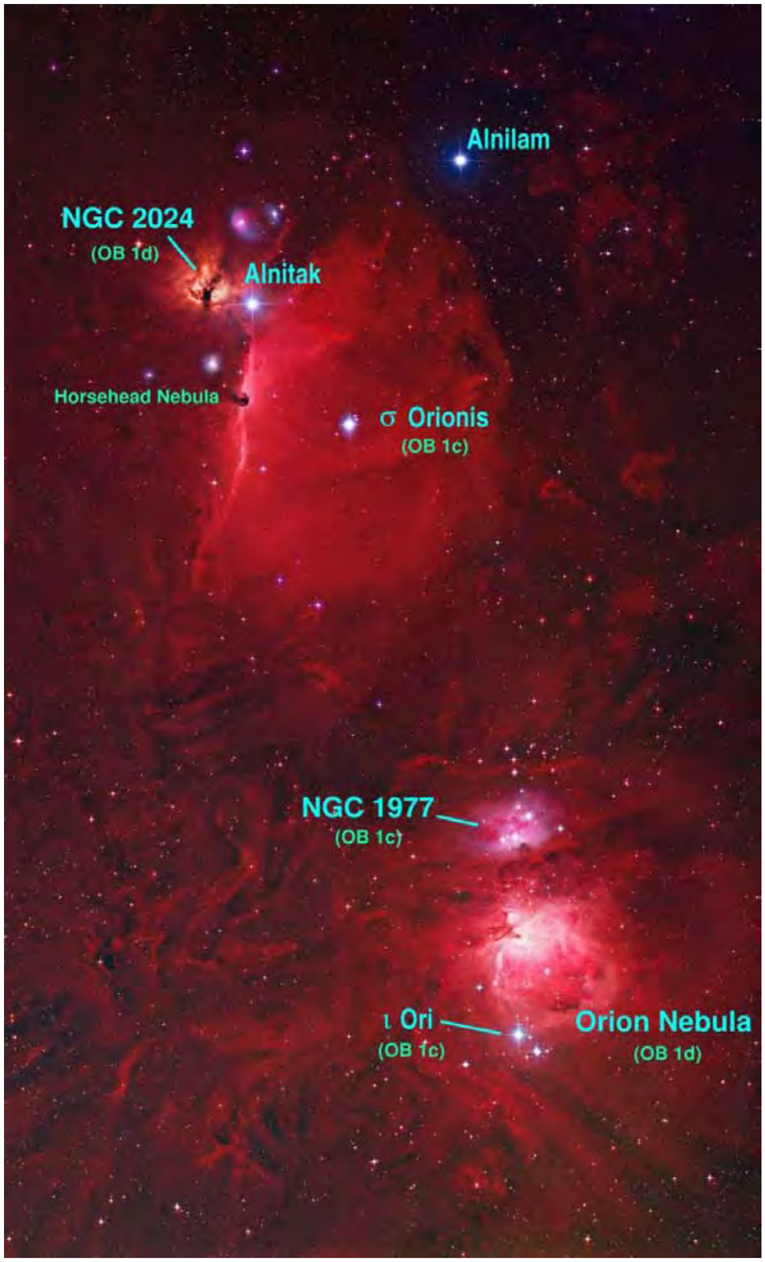


Figure 5.10: The Orion Star Forming Region

The closest star forming region to the Sun which contains massive OB stars. Alnitak and Alnilam are the left and central stars, respectively, in the Belt of Orion. The Orion Nebula and NGC 1977 are two H II regions that make up the Sword of Orion. NGC 2024 is yet another young cluster near Alnitak. The red filamentary structures that run throughout the image is hydrogen gas excited by the UV radiation pumped out by the OB stars. The Horsehead Nebula is part of a molecular cloud which is being eroded away by the B star σ ("sigma") Orionis.

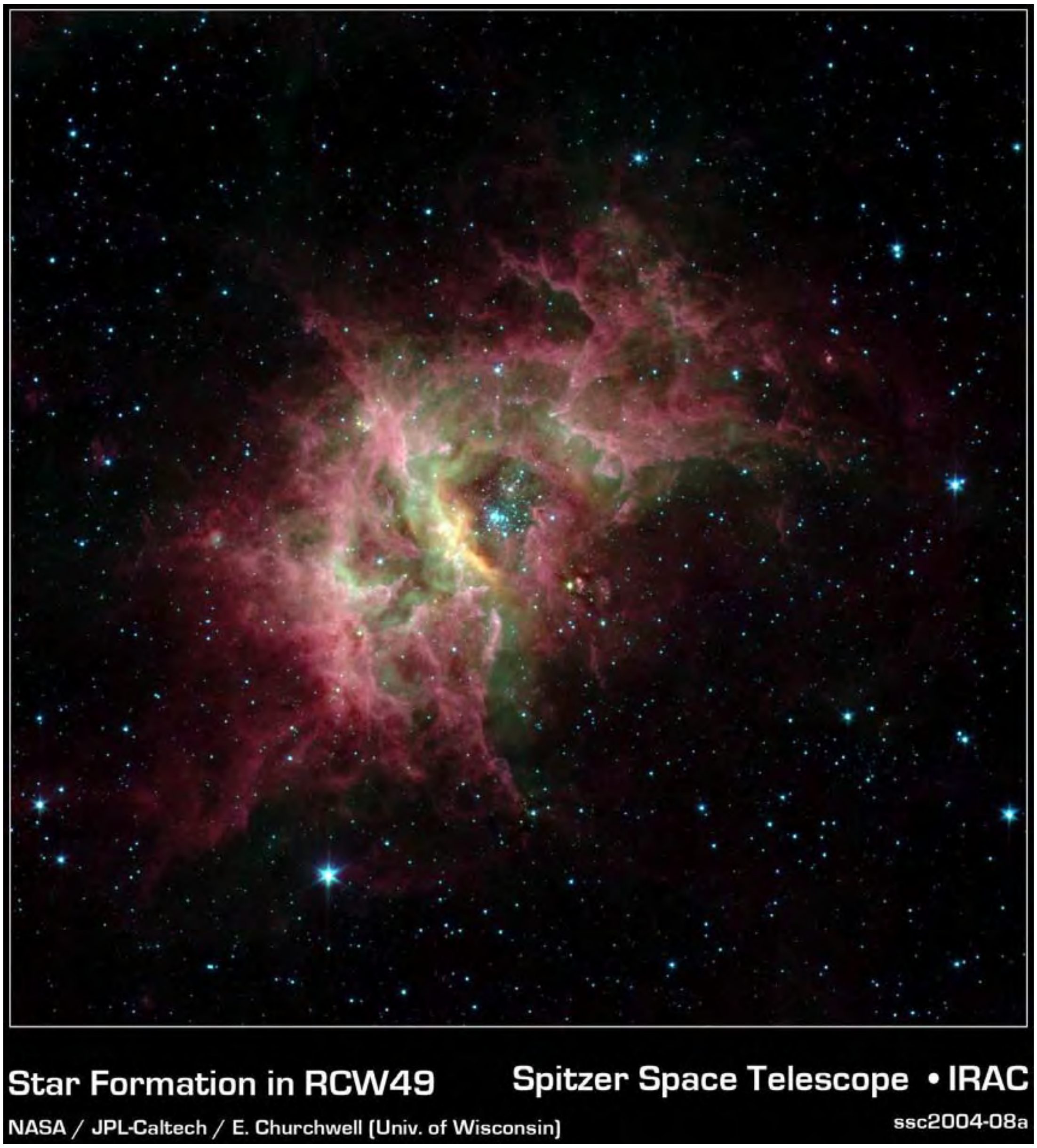


Figure 5.11: RCW 49

A star forming region as seen in the mid-infrared by the Spitzer Space Telescope. Infrared wavelengths of light allow us to peer through some of the thick gas and dust that would normally obscure our view in optical wavelengths. We are looking down into a cavity in the molecular cloud carved out by the young cluster of stars within the H II *region.*

destroying their parent cloud in a weaker, but no less inevitable way.

Star formation can therefore be considered a **self-regulating** process. Long before all of the GMC gas can collapse to form stars, the stars that *have* formed are already acting to destroy the cloud and prevent future star formation in the same GMC. Observations suggest that at most, just 10% of the molecular cloud in a GMC actually ends up in new stars. Over time, the rest of the GMC is completely destroyed, leaving a few **fossil fragments**. The clusters of young stars that remain are no longer tightly gravitationally bound together, and they will slowly wander apart as they orbit the Galaxy.

5.4 A Star's Main Sequence Life

Stars enter their **main sequence** phase once they have attained enough mass to have settled into hydrogen fusion in their cores. Stars will spend roughly 90% most of their lives on the main sequence, converting hydrogen into helium. We saw in §4.1 that the Sun generates its energy via the proton-proton chain. For stars more massive than the Sun, and with core temperatures greater than about 16 million° K, another process becomes important. This is the **CNO cycle** (with the letters standing for the chemical notation for carbon, nitrogen, oxygen). Remembering that a e^+ is a positron, the antimatter equivalent of an electron, and ν is a neutrino, here is a summary of fusion mechanism:

$$\begin{aligned}
{}^{12}\mathrm{C} + {}^{1}\mathrm{H} &\rightarrow {}^{13}\mathrm{N} + \text{energy} \\
{}^{13}\mathrm{N} &\rightarrow {}^{13}\mathrm{C} + e^+ + \nu \\
{}^{13}\mathrm{C} + {}^{1}\mathrm{H} &\rightarrow {}^{14}\mathrm{N} + \text{energy} \\
{}^{14}\mathrm{N} + {}^{1}\mathrm{H} &\rightarrow {}^{15}\mathrm{O} + \text{energy} \\
{}^{15}\mathrm{O} &\rightarrow {}^{15}\mathrm{N} + e^+ + \nu \\
{}^{15}\mathrm{N} + {}^{1}\mathrm{H} &\rightarrow {}^{12}\mathrm{C} + {}^{4}\mathrm{He}.
\end{aligned}$$

The net result in these steps is:

$${}^{12}\mathrm{C} + 4({}^{1}\mathrm{H}) \rightarrow {}^{12}\mathrm{C} + {}^{4}\mathrm{He}.$$

The carbon atom ${}^{12}\mathrm{C}$ is at the start of the reaction, but also winds up at the end, after four hydrogen atoms have been converted into a helium-4 atom. The carbon therefore acts as a *catalyst*, allowing the reaction to proceed, but is not used up or destroyed.

If you start off with a ball of gas with a mass equivalent to that of a star, the self gravity of the gas would cause it to contract. At the beginning, there is nothing to stop the contraction, so the pressures and temperatures in the interior of the star would grow, with the pressure and temperature greatest at the center at the core. At some point, the core temperature and pressure would be high enough to lead to

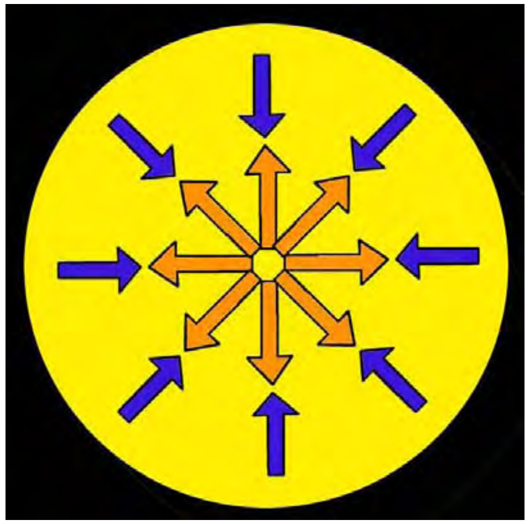


Figure 5.12: Hydrostatic Equilibrium

Throughout most of a star's normal life, the force of gravity of its stellar envelope is balanced by the pressure due to its energy generation at the core. This hydrostatic equilibrium is maintained as long as there is fuel in the core to undergo nuclear burning.

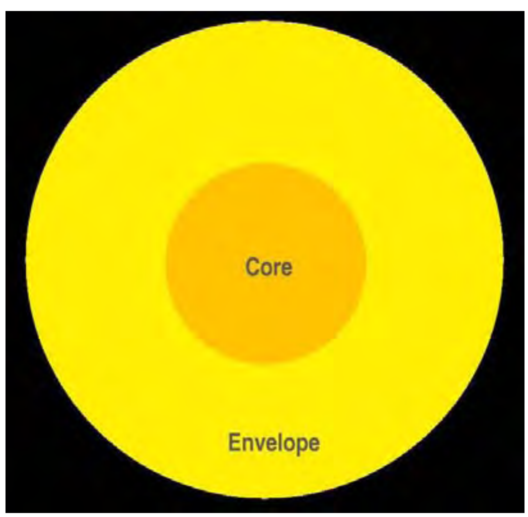


Figure 5.13: A Main Sequence Star

The star during its main sequence *phase has a hydrogen rich core that is slowly converted into helium through nuclear fusion. The envelope is neither dense enough nor hot enough for fusion to take place.*

nuclear reactions starting. The fusion of hydrogen into helium (regardless of whether it is the proton-proton chain or the CNO cycle) generates energy which will tend to leak out to the surface. The energy generated will also pump up the pressure of the interior layers of the star. *This pressure will act to counterbalance the weight of the gas layers which tend to push the gas inward toward the core.* The star is said to be in **hydrostatic equilibrium**. The gravitational forces that squeeze the star together are balanced by the pressure forces that try to push it apart.

Hydrostatic equilibrium is a very important process that keeps the star stable throughout its main sequence life. If the energy generation in the core drops slightly, the weight of the stellar layers above the core will press inward, increasing the center temperature and thereby increasing the rate of energy generation. The resultant extra core pressure pushes out against the outer layers, thereby relieving the weight on the core, decreasing the temperature, and reducing the fusion rate. Gravity and pressure mutually counteract and self-regulate each other during most of the star's life.

5.5 Post Main Sequence Low Mass Stars

However the situation described above cannot continue forever. The supply of hydrogen in the core is not unlimited. At some point, all of the fuel is used up and we are left with a core that is composed of helium. For a Sun-like star, as the hydrogen fusion shuts down, the pressure support which relies on energy generation goes away. There is nothing to hold up the core gas, so it begins to collapse inward. The collapsing core compresses the gas, and heats it up.

During this time, the center of the star consists of a non-burning[5] helium "ash." Just outside the core, there is a layer of hydrogen which, once the weight of the outer envelope presses inward, becomes hot enough to start fusing. This is the **hydrogen burning shell**. The energy generated from this layer is enough to puff up the envelope of gas just above it, causing the outer part of the star to expand. However the amount of energy generated in this shell is not as great as that which was produced in the core. As a result, the total amount of energy that diffuses outward is far less than during the regular core hydrogen burning; the temperatures of the star drops throughout. Recall in the discussion of electromagnetic radiation in §C.3 that cooler objects will tend to glow redder, as their blackbody spectrum shifts to longer wavelengths. Therefore not only does the star expand outward, but it also turns much redder. The star has now evolved into a **red giant**. For the Sun, this occurs after spending 9 billion years on the main sequence. Over a period of 100 million years, the Sun will balloon up until it is 100 times greater than its original size, making it as large as Mercury's orbit, while its luminosity increases by a factor of several hundred. During the red giant phase, winds from the star expel significant material away from the stellar surface,

[5]This is *burning* used to mean a nuclear fusion reaction. This is in no ways similar to the regular fire burning we are familiar with, which is a *chemical* reaction that requires oxygen.

enough to significantly change the mass of the star.

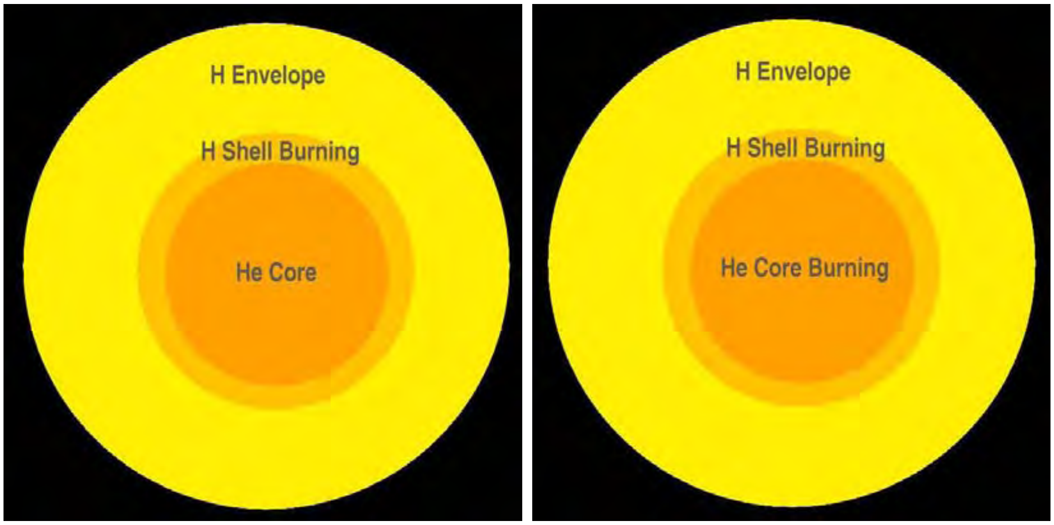


Figure 5.14: Heliuim Shell Burning

Once the core is filled with helium, hydrogen burning will stop and energy production will end temporarily. The star contracts until helium burning can start in a shell around the helium core. This energy production closer to the surface results in a "puffed up" outer envelope which leads to a red giant star. As the core continues to contract, the triple-alpha process that converts helium into carbon begins once the temperatures hit 100 million ° K.

The core is still tiny, only a few times larger than the Earth. It however does continue to contract up to 15 times smaller than its main sequence size, until its temperature reaches 100 million ° K, when another nuclear reaction becomes possible. This is the **triple alpha** process which turns helium-4 into carbon-12 via beryllium-8:

$$\begin{aligned} {}^{4}\text{He} + {}^{4}\text{He} &\rightarrow {}^{8}\text{Be} + \text{energy} \\ {}^{8}\text{Be} + {}^{4}\text{He} &\rightarrow {}^{12}\text{C} + \text{energy}. \end{aligned}$$

The process is so named, because another term for helium nuclei is **alpha particles**. In this reaction, three helium nuclei are fused together to create a single carbon nucleus. Once helium burning starts, the core settles down; the star's luminosity decreases slightly, although its surface temperature grows. This next stage is known as the **horizontal branch**.

The helium burning continues for about 10 million years. Just like the helium ash from the previous stage, carbon ash will build up. Over time, the star will have a carbon core, with a helium-burning layer above it, and a hydrogen-burning shell above that. The process of core contraction as the nuclear fires shut down in the center but fusion beginning again in the intermediate layers result in the star expanding again

Figure 5.15: Comparison Between a Red Giant and Its Main Sequence Precursor

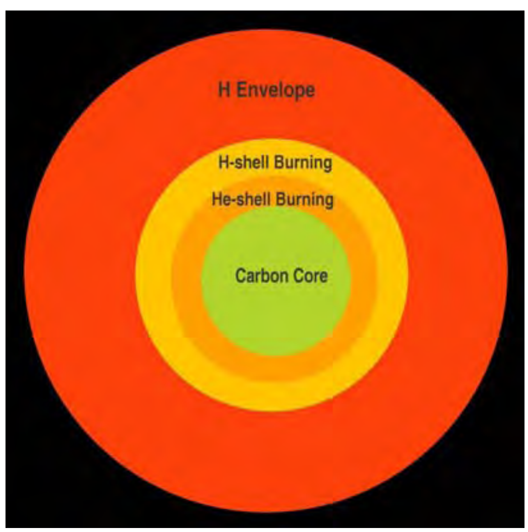


Figure 5.16: A Carbon Core Star

into a **red supergiant**. At this stage, the stellar envelope swells up enormously to become as large as the orbit or Earth or Mars.

After the Sun reaches the red supergiant stage, it is close to the end of its nuclear-burning life. The Sun does not have enough mass for its core to contract to the next set of nuclear reactions, which requires a temperature of 600 million° K. The center continues to cool and contract. The electrons and atomic nuclei get squeezed together tighter, until pressure due to **electron degeneracy** stops the contraction. This is an entirely quantum mechanical effect, but you may think of this as a property of electrons that do not allow them to be compressed past a certain density. The core stops shrinking, and what results is a carbon cinder roughly the size of the Earth, that is incredibly dense (about 10^{10} kilograms per cubic meter) and hot (300 million° K). This is a **white dwarf**, and is the end result of low mass stars like the Sun.

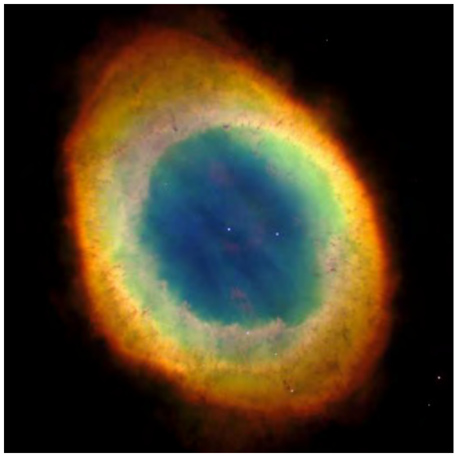
Meanwhile the outer envelope has expanded to as far as an AU from the center. As energy generation stops, the outer layers cool, and the electrons **recombine** back with the atomic nuclei to form ordinary atoms. As the electrons find their way to their ground states (see §C.2), they release photons. The resultant wave of photons push on the gas in the outer envelope, expanding it further and forever out of the gravitational grasp of the shrunken core. This shell of material, a **planetary nebula**[6], grows in size and continues to cool.

A low mass star therefore leaves two relics after its "death." The white dwarf at the center starts out extremely hot (with a surface temperature of 24,000° K) and will continue to glow for billions of years as it slowly cools. Because it is so small, it initially has a luminosity of no more than 0.04 $L_\odot$. Over time, it will gradually turn into a **black dwarf** as it dissipates all of its heat into space. The planetary nebula gas will disperse into the interstellar medium, releasing the heavier-than-hydrogen by-products of nuclear fusion.

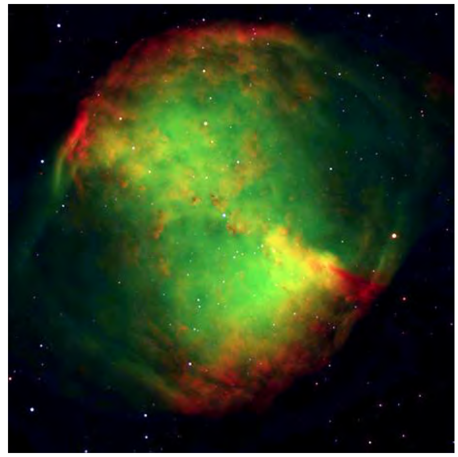
5.6 Stellar End-Life: High Mass Stars

For stars larger than the Sun, their greater mass means that their cores can reach the high temperatures necessary to continue nuclear burning past the helium-fuel phase. Stars that are just 15 $M_\odot$ in mass can have carbon-burning which results in oxygen ash. Higher mass stars can have reactions that create even heavier elements. A cutaway of a high mass star at the end of its life will look like an onion: just underneath its non-burning hydrogen outer surface will be layers of fusing hydrogen, helium, carbon, oxygen, neon, magnesium, and silicon. As each element in the core is used up, nuclear reactions cease, the core contracts, and the temperature and pressure increases enough to start the next wave of fusion. The ash from one nuclear reaction becomes the fuel for the next.

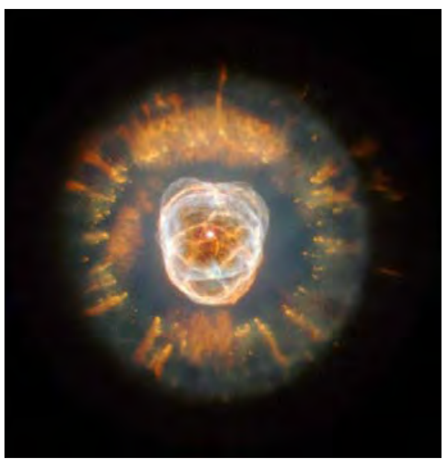
[6]The name refers to their disk-like appearance in primitive telescopes, making them look planet-like.



(a) Ring Nebula



(b) NGC 6853



(c) Eskimo Nebula

(d) IC 418

(e) NGC 6543

(f) M2-9

Figure 5.17: Planetary Nebulae

However the energy generation is less from each new wave of nuclear core burning. Each step staves off the final death of a star by increasingly shorter time steps. For a star that has a mass of 20 $M_{\odot}$, the hydrogen-burning phase lasts for 20 million years, the helium-burning lasts for 1 million years, the carbon for 1000 years, the oxygen for 1 year, and the silicon for 1 week.

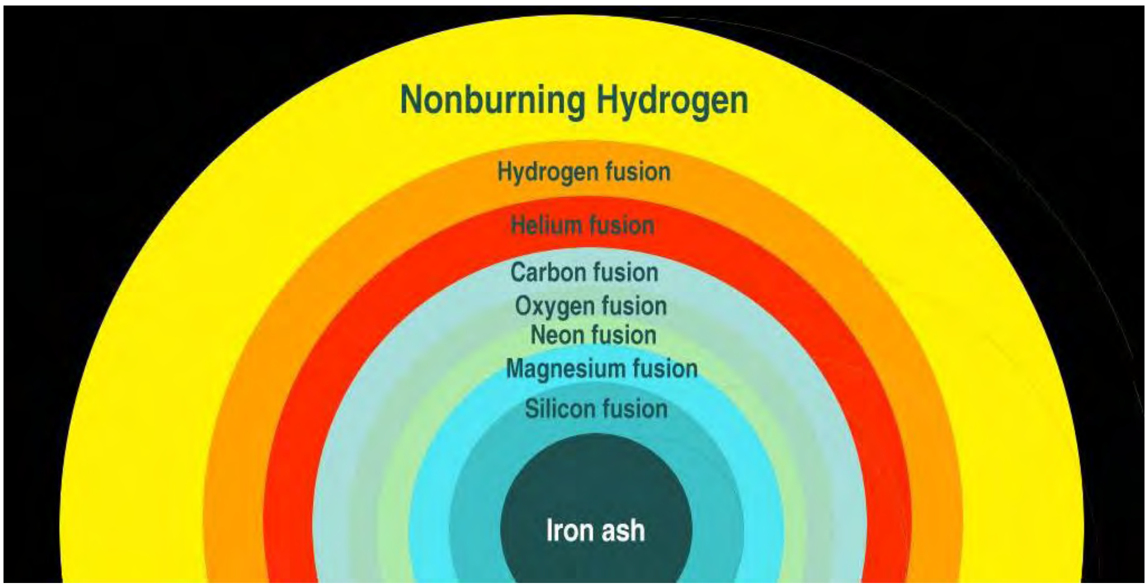


Figure 5.18: The Onion Star
A cartoon showing a cutaway of a massive star nearing the end of its life. Multiple waves of different nuclear burning have resulted in multiple layers of elements, each of which is undergoing fusion. Only the iron ash in the core resists any nuclear reactions.

Eventually an **iron core** will build up from the silicon burning. Iron is the end of the nuclear fusion line, since its nucleus is so compact and dense that there is no way to get any more energy out of it with any type of nuclear reaction. Once the core fills up with iron, the reactions stop, and the star will begin its final collapse. Without any nuclear reactions at the core, hydrostatic equilibrium is gone and the entire star implodes upon itself. When the core reaches a temperature of 10 billion° K, photons energetic enough to split apart atomic nuclei are generated. The last 10–20 million years of fusion effort are destroyed in less than a second as the photons rip apart all of the infalling nuclei until there is nothing left but protons and neutrons, amid a sea of electrons and photons.

As the core collapse progresses even further, the weight of the massive star is great enough to overwhelm the electron degeneracy pressure. At a density of 10^{12} kilograms per cubic meter, the electrons are squeezed so close that they combine with the protons to form neutrons and neutrinos, via a process called **neutronization**:

$$p + e^- \rightarrow n + \nu.$$

Neutrinos (ν) are light particles that almost never react with matter. They fly out of the core as if it was not there. The core consists almost exclusively of neutrons now. Once the collapsing core reaches a density of 10^{15} kilograms per cubic meter, **neutron degeneracy pressure** becomes important enough to halt the collapse. This is again a quantum mechanical effect, like electron degeneracy pressure, and it stops the collapse dead in its tracks, although part of the core may however reach densities as high as 10^{17} or 10^{18} kilograms per cubic meter before rebounding. The time between the start of collapse and the **core bounce** takes less than a second. The rebound sends out a shockwave that passes through all of the outer layers of the star that are *still collapsing* toward the center. The resulting effect is akin to all of the star's outer layers bouncing off of the rebounding core, and getting blown backward in a titanic **supernova explosion**. The amount of energy released is enormous—about 10^{43} Joules—enough that for a short amount of time, the supernova might outshine the rest of the stars in the Galaxy.

A supernova leaves behind several important residues and relics. A **neutron star**, obviously named for its makeup of pure neutrons, is what is left of the core of the massive star. Perhaps even more important for us, the supernova explosion actually *creates* heavy elements. Neutrons expelled from the turbulence of the explosion along with helium nuclei are captured by heavier nuclei to form the rest of the elements in the periodic table that we have yet to mention in this course. Thus *all of the heavy elements that make up the terrestrial planets and which are important for life on Earth were created inside stars or from the explosive aftermath of stars.* The expanding **supernova remnant** propels these elements out into the interstellar medium, where they can eventually wind up in molecular clouds and find their way into the next generation of star (and planet) formation.

Supernova explosions also inject enormous amounts of energy into the interstellar medium. Because massive stars are short-lived—less than 30 million years for OB stars—they usually do not have time to wander far from the GMCs that they were born from before they die. When they reach the end of their main sequence lives, they stir up their environment even more, further disrupting the surrounding molecular gas. In addition to breaking apart gas clouds, shockwaves from supernovae may also have the opposite of *compressing* molecular gas to initiate further collapse of additional cloud cores into future clusters of stars.

One final note about the end lives of stars: nuclear physics tells us that a white dwarf cannot be larger than 1.4 $M_\odot$, beyond which electron degeneracy pressure would not be able to support it. Similarly neutron stars probably cannot be larger than 3 $M_\odot$ for neutron degeneracy pressure to support it. What happens if a star has a core greater than 3 $M_\odot$? If a core of that size collapses, there is *no known physical mechanism to stop the collapse.* The collapse continues forever until the star shrinks to a point, with its gravitational field increasing to where not even light can escape. This is a **black hole**.

Figure 5.19: Crab Nebula

Remnant of a supernova which exploded in our skies in 1054 AD. This image was taken from the Very Large Telescope, with the red filaments representing hydrogen emission, and the blue emission representing energetic electrons in the expanding plasma.

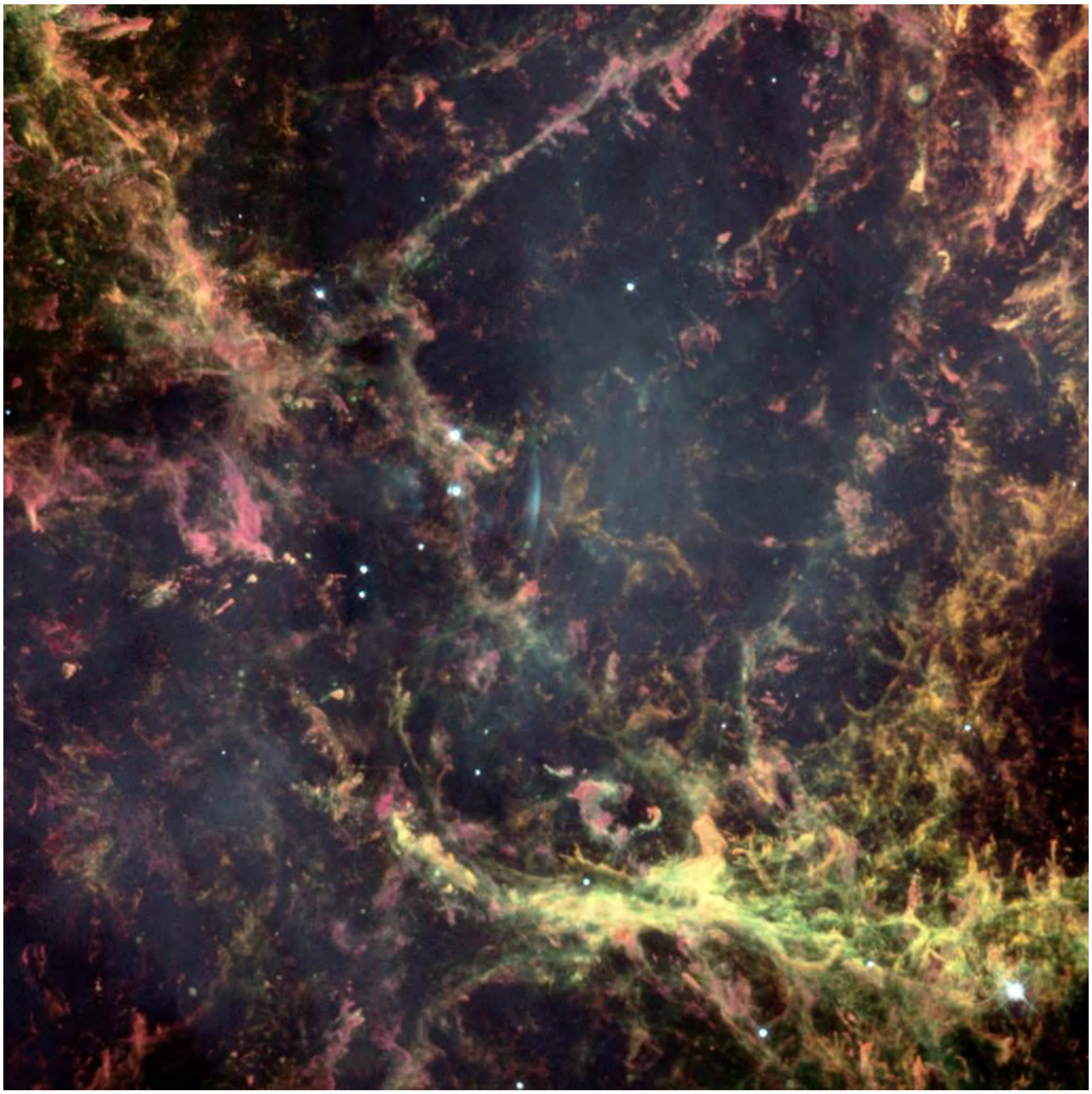


Figure 5.20: Heart of the Crab Nebula

The center of the Crab Nebula as seen by the Hubble Space Telescope, showing the complex array of gas filaments that are speeding away from the explosion center at more than 1000 kilometers per second. The neutron star that is all that is left of the original star is the lower of two bright stars just above the center of the image.

Figure 5.21: Black Hole
An artist's depiction of a black hole orbiting another star. This system originally started out as a binary star pair; one star was much more massive than the other and eventually evolved into a black hole.

5.7 Structure of the Milky Way

We have been discussing the makeup of stars in our Galaxy, but what exactly is the Galaxy? To anyone living at a time before artificial lighting dominated our evening skies, a thin faint band of emission could be seen running across the sky. This has come to be known to us as the **Milky Way**. When early telescopes were turned upon this faint trail, countless more stars not seen by the naked eye suddenly became visible. The Milky Way thus appeared to be a slab or disk containing millions if not billions of stars, with the Sun located somewhere within it.

18th century thinkers and astronomers like Thomas Wright (1711–1786) and Sir William Herschel (1738–1822) believed in the slab or disk model, with the Sun located at the center. When observers tried to count stars along the Milky Way, they were not able to discern any direction that appeared denser than any other. Placing the Sun at the center of the Galaxy therefore was consistent with the evidence of the time.

The true location of our Sun was unresolved until the 20th century, when Harlow Shapley observed globular clusters and a type of **variable stars** within them. From their directions and inferred distances, he was able to show that the globular clusters formed a spherical distribution centered not on the Sun, but at a point nearly 50,000 light years away. (Today we know Shapley had over-estimated his distance by a factor of two.) Thus the Copernican revolution continued: not only was the Earth not at the center of the universe, but neither was the Sun at the center of the Milky Way

Galaxy.

Also during the late 19th and early 20th century, astronomers were studying faint **spiral nebulae** in the sky. There was considerable debate as to whether these were close to the Sun, part of our own Galaxy, or whether they were distant "island universes," separate galaxies of their own. With new and improved methods for determining distances to stars, it became clear over time that these indeed were vast collections of stars that were hundreds of thousands if not millions of light years away. Our Milky Way Galaxy turned out to be another example of these spiral nebulae or **spiral galaxies**.

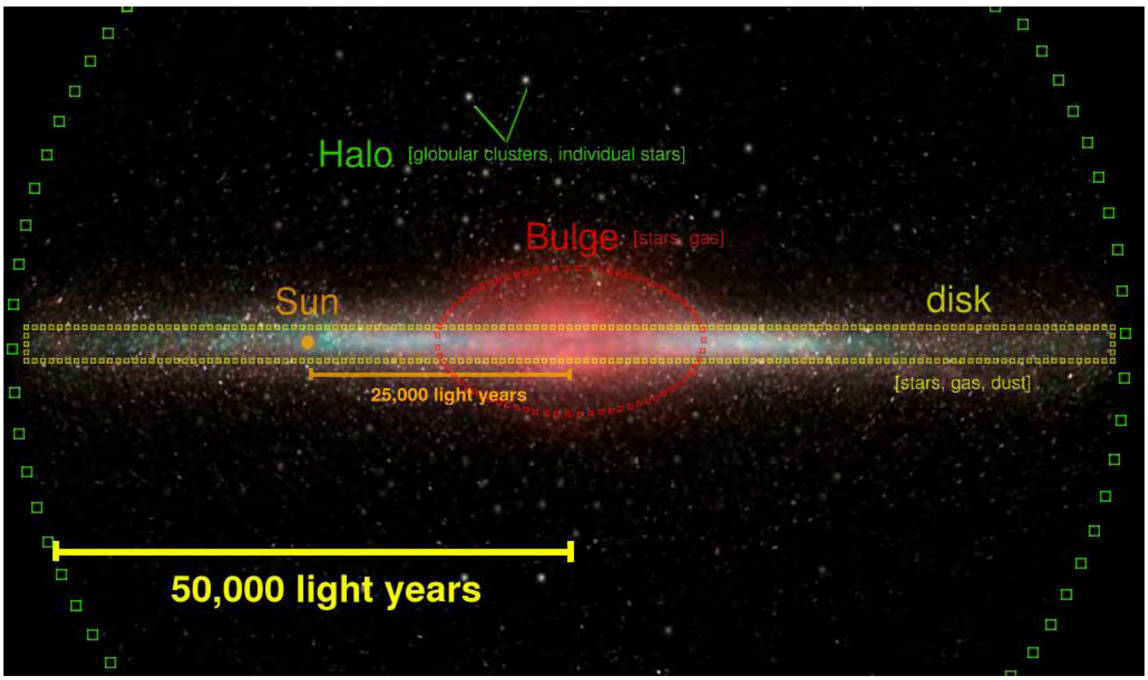


Figure 5.22: The Milky Way

A cartoon of our Galaxy showing all of the major structural components, the disk, the bulge, and the halo, along with rough indicators showing their size. The extent of the halo is actually unknown, and may be many times larger than what is shown.

The modern view of the Milky Way is that of an object consisting of very different populations of stars making up distinct spatial components. The galactic **disk** is a highly flattened structure, about 100,000 light years in diameter, and is relatively thin, "only" about 1000 light years at the location of the Sun (about 25,000 light years from the center). The thickness of the disk is greatest toward the center and decreases outward. The stars here move in roughly circular orbits within the disk around the Galactic center.

The disk also contains the majority of the Milky Way's stars and virtually all of its gas and dust, with a total mass of around 10^{11} $M_{\odot}$. It is in the disk where the spiral arms reside, traced by molecular clouds ranging from small dark globules to GMCs. Because star formation occurs in GMCs, star clusters near the spiral arms tend to be filled with relatively young stars that haven't had time to disperse. The majority of the H II regions bursting out of molecular clouds and illuminated by massive stars are also along the same arms. Supernova explosions from dying OB stars rip through the disk periodically, destroying clouds, creating expanding super-shells of material, and filling the interstellar medium with hot, 1 million° K, x-ray emitting gas. As stars wander away from their birth environments, they disperse into the rest of the disk over time, so that the disk contains a mix of young and old stars. The young stars in the disk is reflected in its color; if we could photograph the Milky Way from the outside, we would notice an overall white-ish color, with a hint of blue in the spiral arms where the massive young stars are located. (We cannot take such an image of the Milky Way, but there are many other galaxies like it that we can look at.)

Toward the center of the Galaxy is the galactic **bulge**. This is a somewhat flattened, elongated, almost football-shaped collection of stars that is aligned with the plane of the galactic disk. Bulge stars further from the center are older, although there is a group of younger stars close to the center. The inner center also contains a ring of gas and dust, with ongoing star formation. Instead of being circular and aligned in the same plane, the orbits of the bulge stars can be highly elliptical, with largely random inclinations. However if all of the bulge stellar motions were averaged together, there would be some net rotation about the Galactic center. Because there is a slightly older population in the bulge (outside of the innermost star-forming regions), the collective color of the stars is yellow-ish white.

Finally around the disk and bulge is the spherical galactic **halo**. The halo contains virtually no gas and dust, and therefore has no star formation (although as we will see in the next section, there is some very tenuous gas in the halo). The halo stars are relatively old, which give it an overall reddish color. The halo is sparsely populated, containing only about 1 billion $M_\odot$ worth of stars. About 1% of the halo stars are found in the 150 **globular clusters**, which contain some of the oldest stars in the Galaxy. These are small densely packed, spherical groupings of 10,000 to one million stars. These like the rest of the halo stars move around the Galactic center in randomly oriented orbits.

5.8 The Evolution of the Milky Way

As we have seen in the preceding sections, the star life cycle increases the overall amount of heavy elements—any of the elements with mass greater than hydrogen and helium. Supernova explosions and stellar winds expel these heavy elements back into the interstellar medium. This enriched gas gets recycled back into subsequent generations of stars. If we can run the a movie showing the evolution of the Milky Way backward, we will see the enrichment of heavy elements in the stars drop as we move further back in time. One possible consequence of fewer heavier elements in the past is that rocky planet formation would have been rarer.

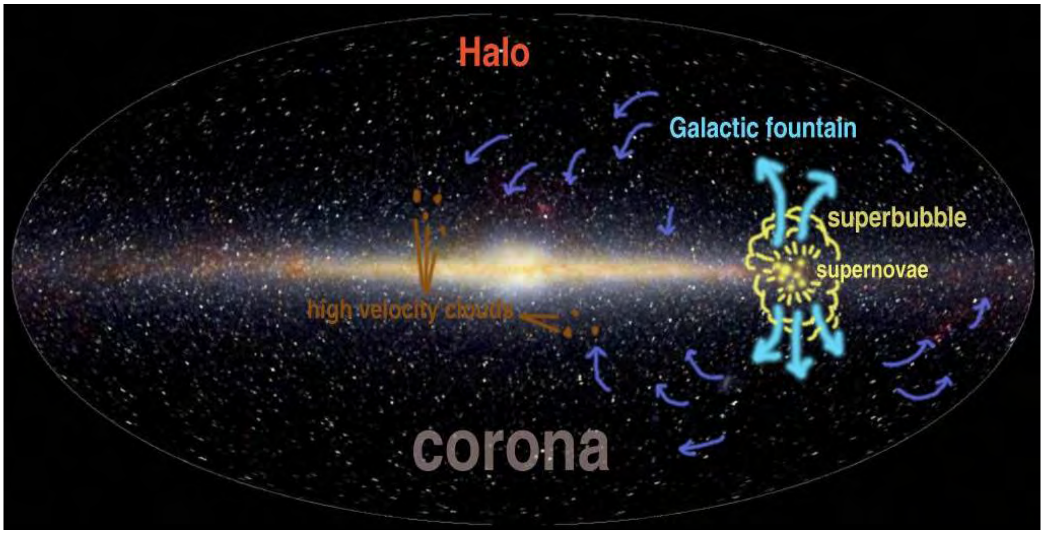


Figure 5.23: The Galactic Fountain
A model of how gas can be ejected out of the disk of the Galaxy. Multiple supernovae explosions create an expanding superbubble which expels hot plasma gas out of the disk. This gas forms a tenuous corona in the halo, and may fall back to the Galactic plane after it cools in the form of high velocity clouds.

The largest supernova explosions not only eject gas into the disk, but they throw up material *out* of the disk. In particularly active star forming regions, multiple OB stars will explode in supernovae within just a few million years of each other. The hot, ionized gas that expands from one one explosion overlaps with the ionized bubble from another. These multiple bubbles add up together to form a **superbubble** which can expand vertically out of the disk of the Galaxy. This hot gas is squirted into the halo around the Milky Way. Eventually the gas is expected to cool and fall back down to the disk. This entire mechanism has been termed the **Galactic fountain**, and it is estimated that about 10 $M_\odot$ worth of material gets recycled in this manner. The hot million degree gas makes up a **gaseous corona** in the halo of our Milky Way. We have not found conclusive proof that this coronal gas eventually falls back into the disk, but clouds of atomic hydrogen falling at high speeds toward the Sun have been observed. Distance determinations to these **high velocity clouds** are difficult to make, but the majority of them appear to be well beyond the stars in the Galactic disk.

When we observe galaxies elsewhere in the universe, we see evidence of past as well as ongoing collisions between them. Astronomers have speculated that this might have happened to the Milky Way as well, and conclusive proof of this came in 1994 when a small galaxy, now called the **Sagittarius dwarf galaxy** was discovered. The

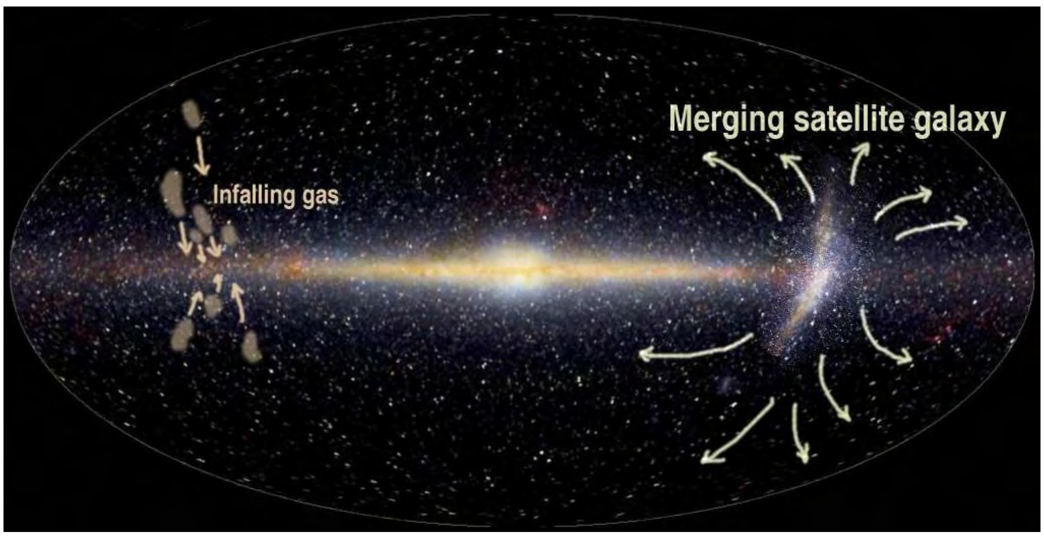


Figure 5.24: Growth of the Milky Way

The two methods by which a spiral galaxy like the Milky Way may grow over time: It can absorb infalling gas directly into the Galactic disk, where it is added to the molecular cloud reservoir available for star formation. Mergers with smaller satellite galaxies result in the smaller galaxy getting pulled apart by the gravity of the Milky Way, with the stars dispersing into the halo.

Sagittarius dwarf is colliding on the far side of the Galaxy, and its stars are very faint, which explains why it had remained hidden for so long. The gravitational interactions from this collision is thought to eject stars into the halo. Several globular clusters that were thought to be members of our Milky Way have turned out to be likely gravitationally-captured objects from the Sagittarius dwarf.

We will discuss the details of the formation of galaxies in the next Lecture. However the interactions of our Milky Way with the Sagittariuis dwarf suggests that galaxies are not static objects, but continue to grow and evolve over time even after their initial formation. Computer simulations have shown that mergers and collisions between galaxies tend to distribute stars from both galaxies into the galactic halo and bulge. Infalling gas however tends to settle into the galactic disk.

A story of how our Milky Way formed and evolved might therefore be as follows: Our Milky Way might have originally started out as an **elliptical galaxy**, which grew from a series of mergers of many smaller galaxies. This initial elliptical galaxy still exists today in the form of the Galactic bulge. In the intergalactic environment that the Milky Way evolved in, there was however much more free gas than there was dwarf galaxies. The Milky Way therefore grew over time by accreting infalling gas, which collected in the Galactic disk, while its halo and bulge have remained relatively small.

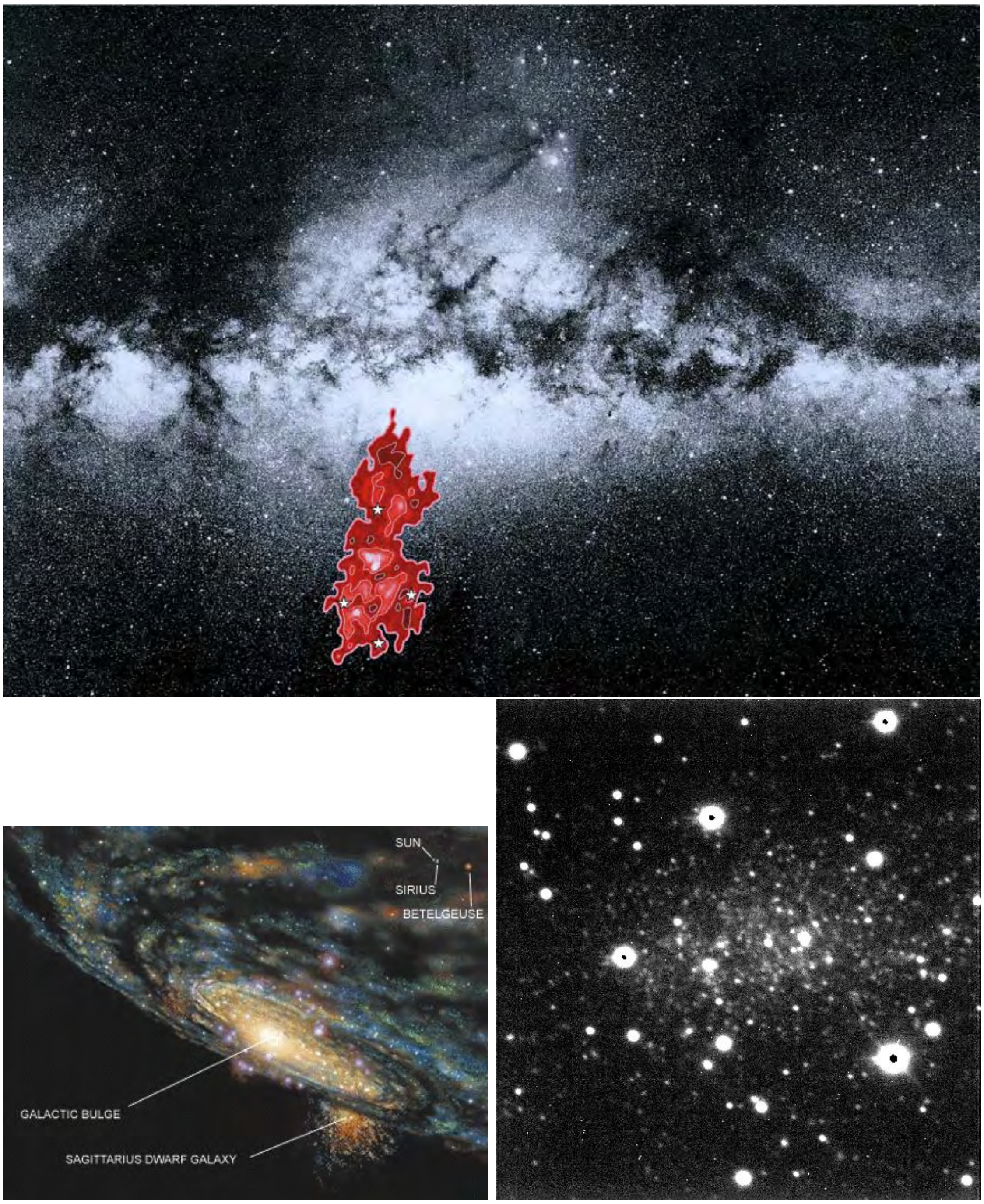


Figure 5.25: Sagittarius Dwarf Galaxy
Top: *Red contours show the location of the Sagittarius Dwarf Galaxy, on the other side of the Galactic Center, after having plunged through the disk of the Milky Way.* Bottom Left: *An artist's depiction of where the Sagittarius Dwarf Galaxy is relative to the Sun.* Bottom Right: *Stars and globular clusters thought to be associated with this satellite.*

Lecture 6

The Rest of the Universe and the Big Bang

6.1 The Extra-Galactic Zoo

Our Milky Way Galaxy turns out to be just one galaxy among hundreds of billions of other *external* galaxies (to differentiate them from our own). Most are **normal galaxies** which have luminosities that do not change over time, and whose light output is due mostly to stars. A second category of extra-galactic[1] objects are the **active galaxies**, which have highly variable luminosities that are nonstellar in origin. (Note that we have been spelling terms like *galaxy* and *galactic* with a capital *G* when referring to our own Milky Way, while using the lower-cased version for any other galaxy.)

6.1.1 Normal Galaxies

In the 1920s, when there was still some debate as to the true nature of the faint, extended nebulae that were being observed by astronomers, it was clear there was enough variety in their appearances that a **morphological classification** was devised. The first scheme was introduced by Edwin Hubble, and we still use his terminology today.

The four basic **Hubble classes** are **elliptical**, **spiral**, **lenticular**, and **irregular** galaxies. The spiral and lenticular classes are divided additionally into **barred** and **unbarred** varieties. Each Hubble class is also subdivided into **Hubble types** using letters and numbers, which refine the classifications further.

The Hubble class of elliptical galaxies are objects that look elliptical in shape (obviously!), with a bright core and a luminosity that drops off smoothly away from the center, with little or no clumpiness. The Hubble types range from E0, for ellipticals that look circular, to E7, for ellipticals that are highly elongated. Note that this

[1]Meaning anything outside of our own Galaxy.

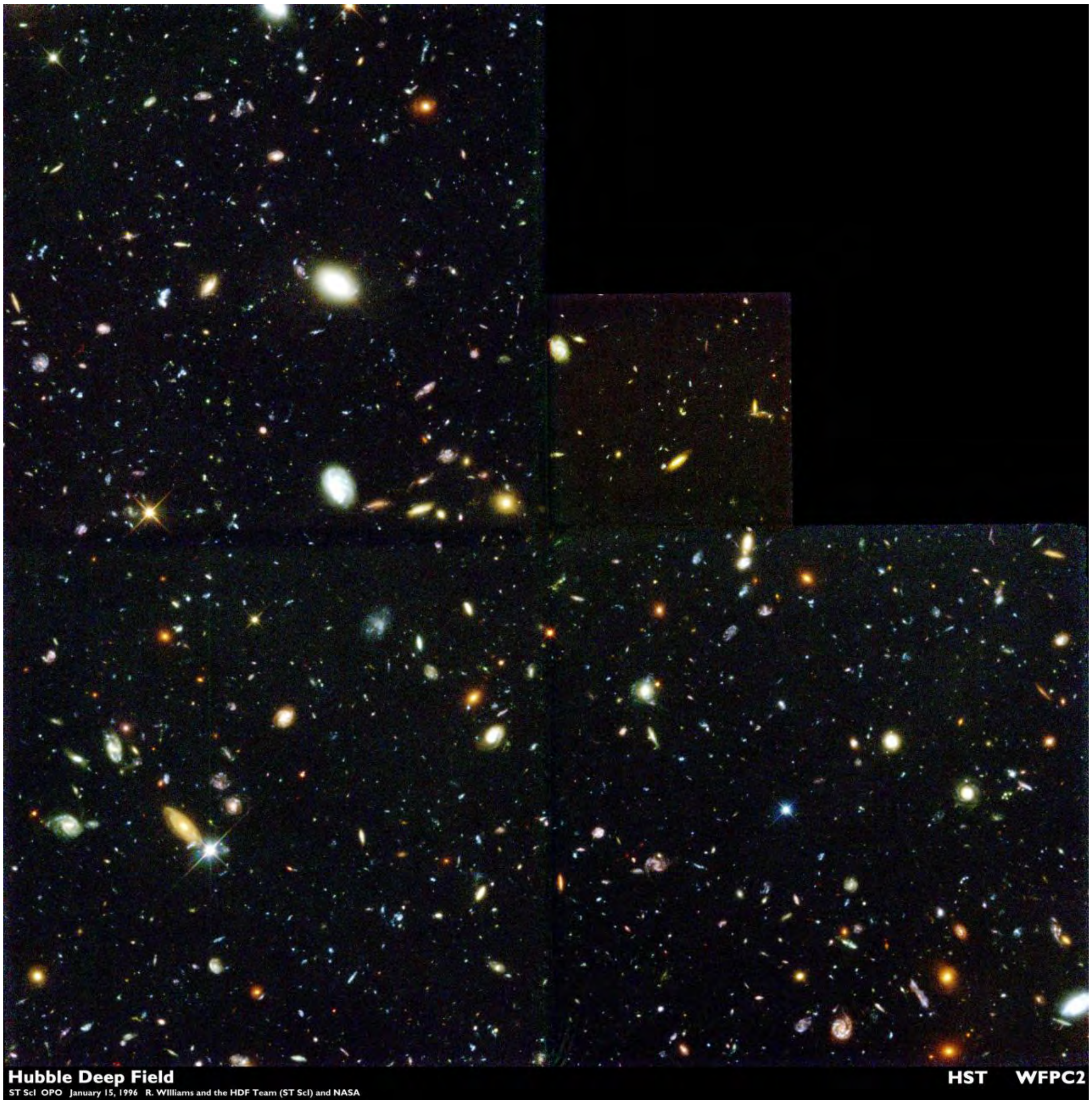


Figure 6.1: Hubble Deep Field

The Hubble Space Telescope was pointed in the same direction toward the handle of the Big Dipper for 10 consecutive days in December 1995. This is a region of the sky with no nearby galaxy clusters. The resultant image from this long integration is shown above. More than 1500 galaxies are revealed, most never before observed by even the largest ground-based telescopes. Views like this reveals that the observable Universe is filled with galaxies.


(a) M89 (E0) (b) M105 (E1)


(c) M32 (E2) (d) M49 (E4)


(e) (E5) (f) (E6)

Figure 6.2: Elliptical Galaxies
Examples of a range of elliptical galaxies with different Hubble types. All of these images were taken using telescopes at the National Optical Astronomy Observatories (NOAO).

classification is made by the elliptical's apparent shape. It has nothing to do with its actual three-dimensional shape, since a long, cigar-shaped object might look circular with the right perspective. Ellipticals are the most common type of galaxies, making up more than 60% of all those that are observed. Ellipticals also come in a range of sizes. The most prominent are the **giant elliptical galaxies**, which can be 3–4 times larger than the Milky Way. Most ellipticals belong to the faint **dwarf elliptical** class, containing only a few million stars. Dwarf galaxies may be undercounted in galaxy surveys because they are more difficult to find than their brighter cousins. The stars in elliptical galaxies tend to be old, with no significant star formation in the last 10 billion years. This makes sense since there is little gas and dust left in them.

The class of spiral galaxies is what our Milky Way falls under. A bright central bulge is surrounded by a flattened stellar disk. The disk further has spiral arms, traced by clumpy dust clouds. Spirals are further categorized by how tightly the spiral arms wrap around the bulge, and whether the galaxy has a bar at the center of the spiral. The Hubble types Sa, Sb, and Sc are for unbarred spirals, with Sa galaxies having the tightest arms and largest bulges, and Sc galaxies having the loosest wound arms and smallest bulges. Because of the relationship to bulge size, the exact Hubble type for a spiral galaxy can be guessed even if the galaxy is nearly edge-on. Barred spirals have the inner parts of the spiral arms meeting at the ends of a stellar bar. The barred spiral types are SBa, SBb, and SBc, with SBa tightly wound, and SBc loosely wound. Barred spirals are only evident if the galaxy is seen face-on. Spiral galaxies are often seen close to edge-on, such as M 31, the Andromeda Galaxy. However many spiral galaxies, including our own Milky Way, show evidence of a bar, and it may turn out that all spirals have bars, with many too small to be seen. And because of the physical and chemical similarities between barred and unbarred spirals, some extra-galactic astronomers[2] do not even bother differentiating between the two categories.

The disks in spirals have copious gas and dust allowing for ongoing star formation. As a result, both young and old stars can be found mixed together in the disk. The lack of gas in the bulge and halo means only older stars are found in those regions.

Lenticular galaxies are similar to spiral galaxies by having a disk and a bulge. However their disks possess no spiral arms. The bulge may also be relatively large, compared to the disk, and may contain a bar-like structure as well.

Spirals and lenticulars make up about 30% of all galaxies, with about 60% of each class having noticeable bars.

Irregular galaxies do not have the clear symmetries or regular shapes found in the

[2] "Extra-galactic" here refer to *outside our own Galaxy*. Extra-galactic astronomers study galaxies and other objects outside of our own Milky Way. Another term you might have heard of is *cosmologist*, which is someone who studies the Universe as a whole. The research associated with cosmologists include the inception of the Universe at the Big Bang along with its subsequent expansion and evolution (see 6.2). Some cosmologists work at such large scales that they consider individual galaxies to be too insignificant to worry about!

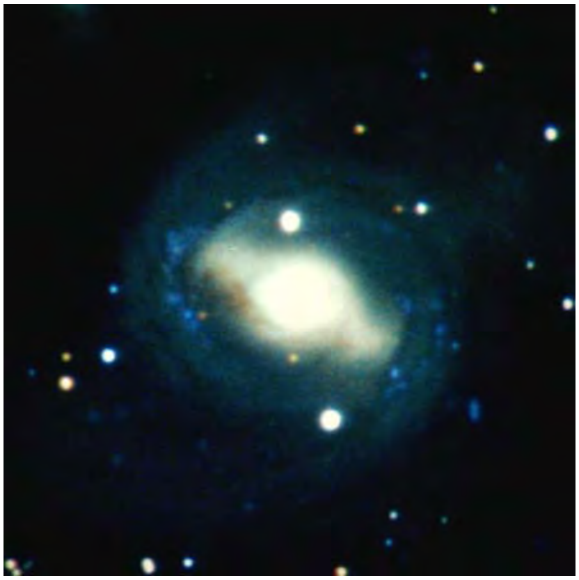
(a) NGC 7217 (Sab)

(b) M77 (Sb)

(c) M99 (Sc)

Figure 6.3: Unbarred Spiral Galaxies

Examples of unbarred spirals taken using NOAO telescopes. The Sab type is intermediate between Sa and Sb. The Hubble classification scheme places spirals with more tightly wrapped arms in the Sa category, and galaxies with the loosest spiral arms in the Sc category.



(a) NGC 4650(SBa)

(b) M91 (SBb)

(c) NGC 1073 (SBc)

Figure 6.4: Barred Spiral Galaxies

Examples of barred spirals taken using NOAO telescopes. The SBa type is the most tightly wrapped barred spiral, while the SBc type is the least tightly wrapped. Note that the spiral arms terminate at the ends of the bar in these galaxies.

(a) M85 (S0)

(b) NGC 936 (SB0)

Figure 6.5: Lenticular Galaxies

Examples of lenticular galaxies taken using NOAO telescopes. Lenticulars can be subdivided into the unbarred (S0) and barred SB0 types.

(a) Large Magellanic Cloud

(b) Small Magellanic Cloud

Figure 6.6: The Magellanic Clouds
The closest irregular galaxies to the Milky Way, taken from NOAO telescopes.

(a) IC 4182 (Irregular I)

(b) NGC 55 (Irregular I)

(c) NGC 3077 (Irregular II)

Figure 6.7: More Irregular Galaxies

previously described Hubble classes. Some have traces of spiral arms or hints of a disk. These are the **Type I** irregulars, which have regions of vigorous star formation, with vast H II regions and brilliant OB stars. The Magellanic Clouds are the two closest examples of this sub-class. **Type II** irregulars are much more chaotic looking, often containing odd arms, loops, or explosive filaments. These may be the result of mergers or collisions between neighboring galaxies. Irregulars are the least common of the traditional galaxy classes, making up about 15% of all observed galaxies. Finally there is a small fraction of **peculiar** galaxies as well, that can not be made to easily fit any of the Hubble categories.

Edwin Hubble plotted the different Hubble types in his classification scheme into a "tuning fork" diagram, with the ellipticals at the "handle", and the barred and unbarred spirals making up the "tines" of the fork. One can look at this arrangement of galaxies and think that it also implied an evolutionary sequence: galaxies started in a spiral class, and over time "evolved" into the elliptical class, as all of the gas and dust was used up in star formation. Although it is tempting to suggest that galaxy morphologies can be so easily explained, there has been little evidence to support such evolutionary trends for galaxies.

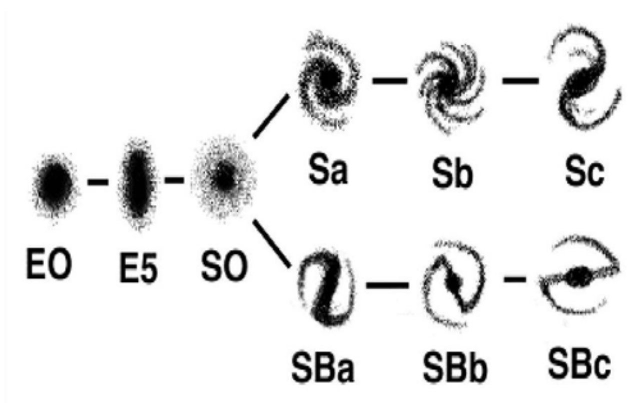


Figure 6.8: Hubble's "Tuning Fork" Classification
Edwin Hubble's classification scheme for galaxies involved placing spirals and barred spirals at one end, and elliptical galaxies at the other, with lenticulars between the "handle" and "tines" of a tuning fork. The various types of spiral galaxies are further divided by how tightly the spiral arms wrap around. The location of an elliptical along the handle depended on how much it deviated from a circular shape.

6.1.2 Active Galaxies

Many different categories of active galaxies have been found, with a confusing mix of names and sub-categories: Seyfert 1 galaxies, Seyfert 2 galaxies, quasars, radio galaxies, Fanaroff-Riley type 1 and type 2 galaxies, blazars, BL Lacertae objects, LINERs. All these types were discovered separately, with their own peculiar observational characteristic that made them distinct. However despite these differences, it is now clear that all of them release an enormous amount of energy from a very small volume at their centers. This radiation is usually in far excess of the light from the galaxy's stars, at least 10^{10} times or more the power of the Sun. This light is also *nonstellar*, meaning it has spectral characteristics that are completely different than the blackbody radiation emitted from stars.

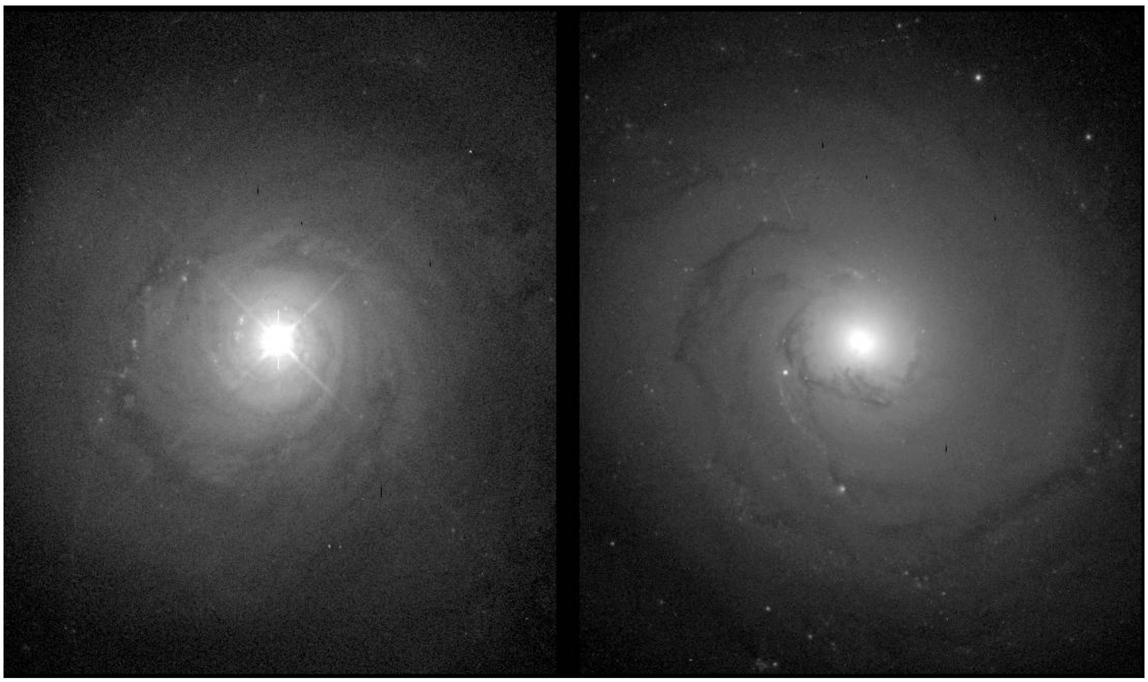


Figure 6.9: Comparison of Active and Normal Galaxies

Hubble Space Telescope images showing NGC 5548, an active galaxy (specifically a "Type 1 Seyfert;" on the left); and a normal galaxy NGC 3277 (on the right). Both have stellar disks containing spiral arms. However the point-like nucleus of the active galaxy is so bright that it has saturated the detector and has produced the "diffraction spikes" artifacts pointed diagonally away from the center. (Diffraction spikes are artifacts due to the wave nature of the incoming light, as the light bends around the secondary support structures inside a telescope. They are always present, but are apparent only around extremely bright point sources.)

The energy output originates from a compact **active galactic nucleus** or **AGN**. AGNs appear as points in optical images of their host galaxies. The best optical and radio imaging measurements show them to be so extraordinarily compact, that they are *unresolvable*–they are so small that no detail or sub-component can be seen.

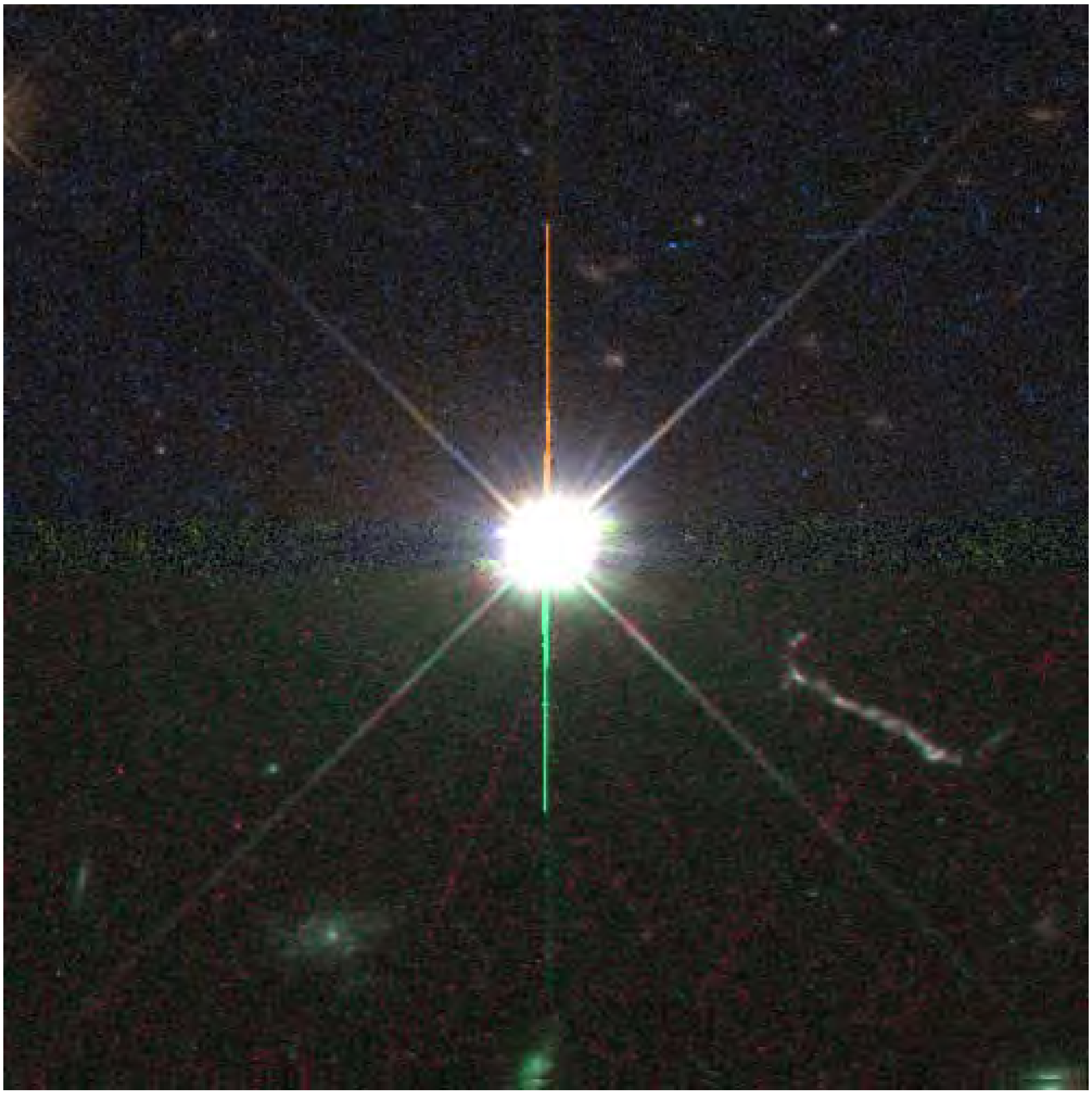


Figure 6.10: 3C273

The brightest quasar in the sky as seen by the Hubble Space Telescope. The nucleus of the quasar shines as a brilliant point source (which results in the "diffraction spikes" imaging artifacts emanating from the center). A jet pointing back to the nucleus can be seen to the lower right, as well as numerous fainter, foreground galaxies.

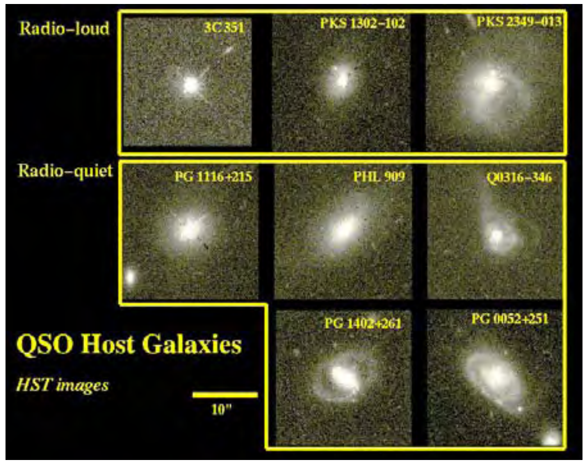


Figure 6.11: Quasar Host Galaxies

Hubble Space Telescope observations of quasars showing the faint host galaxies which surround the bright active nuclei.

Another way to measure them is by analyzing their variability in energy output which can be as short as a few hours. This implies that the luminous source must be very small, less than a few light years across, for it to change brightness so quickly.

Quasars are some of the most distant and powerful of AGNs.[3] Some like 3C 273 show faint jets of material emerging from the nucleus. Jets are also visible in the much closer *radio galaxies*. One example, Cygnus A has two radio-bright "lobes" of gas, connected to either side of the nucleus by a pair of narrow jets.

The only type of astronomical object that are powerful enough to drive AGNs, as well as to explain some of their properties, is thought to be **supermassive black holes**. As they accrete gas and stars, their gravitational energy is converted into radiation. Gas clouds falling toward the central black hole will collect into an **accretion disk**. The disk heats up, due to the viscosity of the gas as it spirals in. The heating effect increases with the strength of the gravitational field, so the inner edges of accretion disks will reach the highest temperatures. The inner accretion disk will therefore emit brightly in electromagnetic radiation. As matter falls into a black hole, about 10% of the rest mass energy[4] can be converted into radiation. This energy is far more than can be radiated by an equivalent mass of stars in a star cluster, and is perhaps the most efficient way to generate energy from matter outside of annihilating

[3]The name originates from the term *quasi-stellar radio source*. Quasars were first detected at radio wavelengths. However when astronomers looked at them at visible wavelengths, they could only see a star-like point source. Spectroscopic measurements showed them to be much too far away, and hence too bright, to be actual stars.

[4]We can use Einstein's mass-energy equivalence formula $E = mc^2$ to determine how much energy can be converted from a lump of ordinary matter. In the formula, m is the mass of the matter, c is the speed of light, and E is the total energy within that matter.

it with its antimatter counterpart.

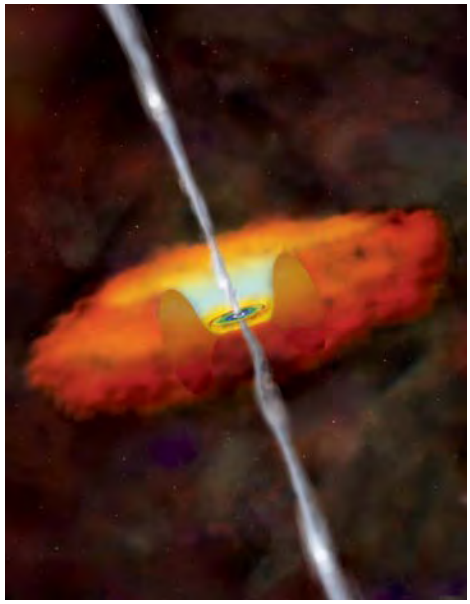


Figure 6.12: Model of an AGN
The standard model of what powers active galactic nuclei: a supermassive black hole sits at the center of the host galaxy. The black hole accretes from the surrounding gas and dust, and generates jets that eject gas from the "north" and "south" poles of the black hole. The painting shows a cutaway through the surrounding torus of gas so that we can see the central black hole.

Many hypotheses have been proposed to explain jets from AGNs. It is assumed that the jets point at right angles to the accretion disk, i.e., aligned with the rotation of the disk. One theory suggests that the accretion disk may thicken and bunch up at its inner edge. The radiation pressure from the high temperature gas at this region may be enough to accelerate and funnel some fraction of the gas into jets.

6.1.3 Galaxy Clusters and Superclusters

Our own Milky Way is in a small cluster of galaxies, the Local Group. This is a gravitationally bound grouping of galaxies, within a volume of space 6 million light years across. Each member of the group moves according to the net gravitational force of the entire Local Group, and galaxies cannot escape unless they are ejected from a collision or other close encounter. The two largest galaxies are spirals, the Milky Way and M 31 (the Andromeda Galaxy). The next largest are the two Magellanic Clouds and the small spiral M 33 (the Triangulum galaxy). The Milky Way and Andromeda each have a small flock of smaller dwarf elliptical and irregular galaxies that orbit nearby. A current census of the Local Group shows 30 members, but because dwarf elliptical members are difficult to find, we may expect the membership to grow as our telescopic instrumentation improves.

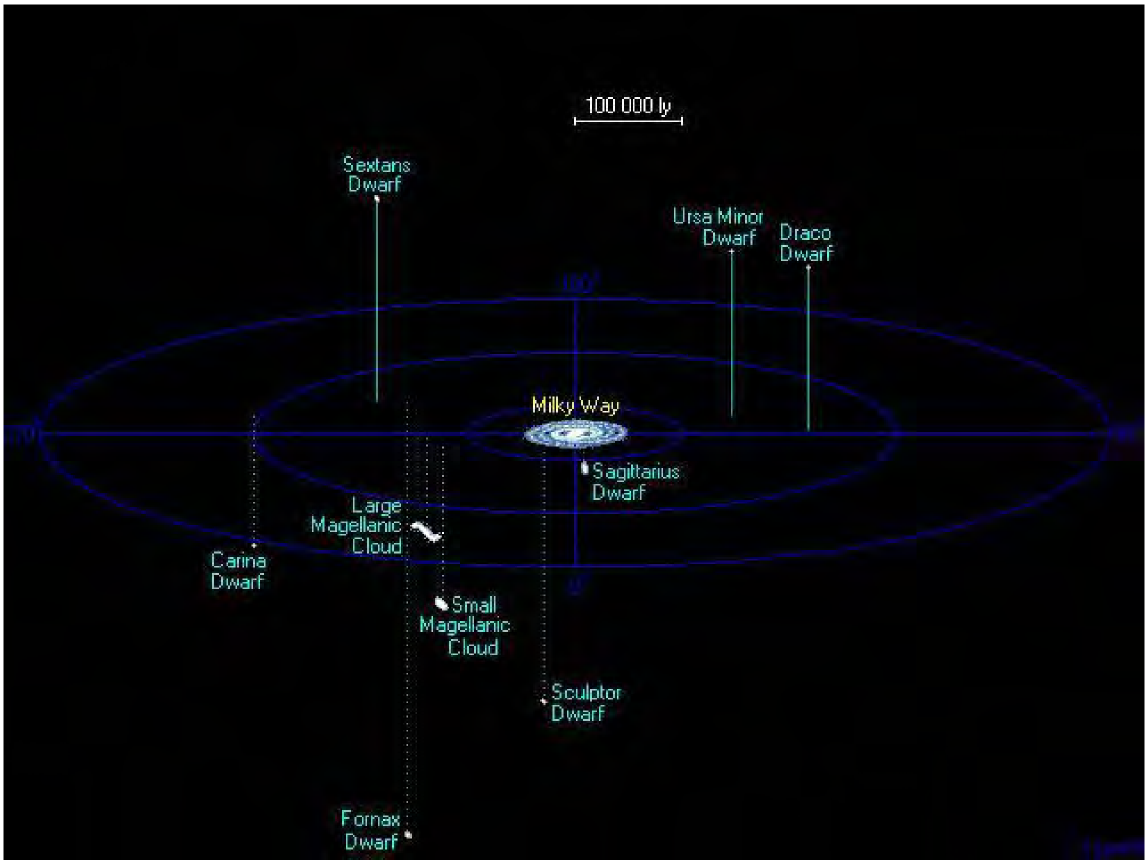


Figure 6.13: Satellite Galaxies of the Milky Way
The immediate vicinity of our Milky Way galaxy showing our satellite galaxy neighbors.

As clusters go, the Local Group is rather sparse and small. Other clusters are far **richer**, meaning not only do they have larger memberships (up to thousands of observable galaxies), but they are also more densely packed. In fact the typical rich cluster is comparable in diameter to the Local Group, but there are far more galaxies at the cluster cores. Two well-known clusters are the **Coma cluster** and the **Virgo cluster**. Each has over a thousand known members.[5] The Coma cluster is about 300 million light years from the Milky Way, and is a spherically shaped cluster consisting mainly of elliptical and lenticular galaxies. The Virgo cluster is closer, at 60 million light years, and is irregular in shape, with a mix of spirals as well as

[5]Remember that these are *observed* galaxies. Many galaxies, such as dwarf ellipticals, are below the current observable threshold, and therefore would not be counted.

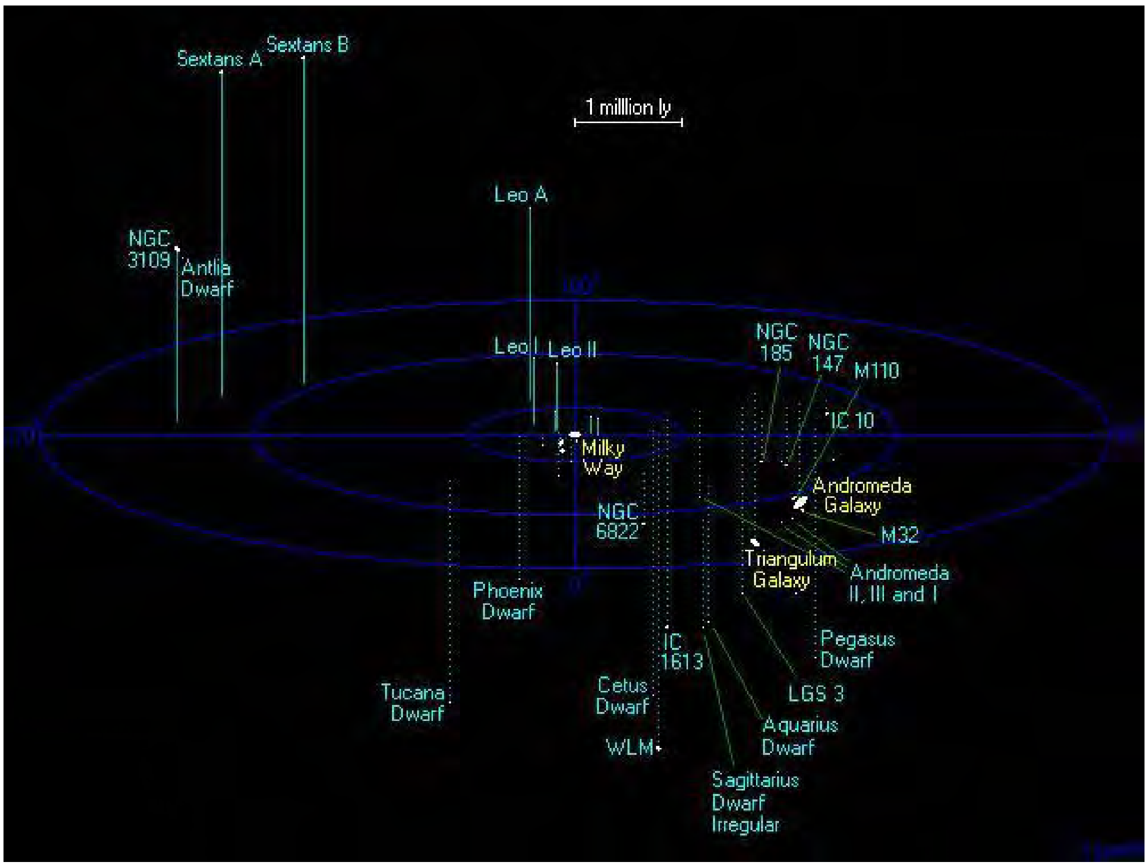


Figure 6.14: The Local Group

Within 5 million light years of the Milky Way are a loose collection of galaxies that make up the Local Group. The two massive galaxies, our own Galaxy and the Andromeda Galaxy (M31), dominate the cluster gravitationally.

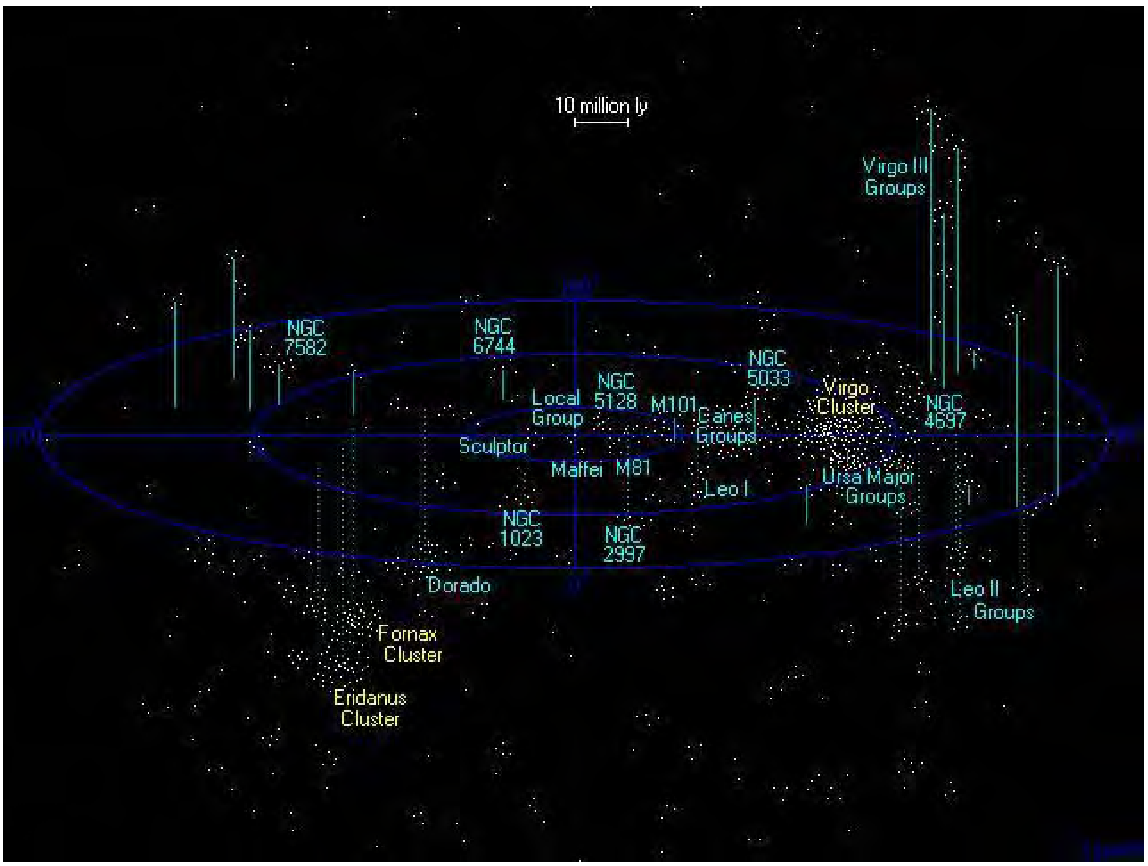


Figure 6.15: The Virgo Supercluster

A plot showing galaxy clusters within 100 million light years of the Milky Way. This volume of space is dominated gravitationally by the Virgo Supercluster. Although our Local Group is speeding away from it currently, it will probably slow down, reverse course, and eventually fall toward the local supercluster and merge with it in the far distant future.

Figure 6.16: The Virgo Cluster

The nearest supercluster is located in the constellation of Virgo. Although the center of the Virgo cluster has a few giant ellipticals (such as M87 and M86), there are also many spirals as well.

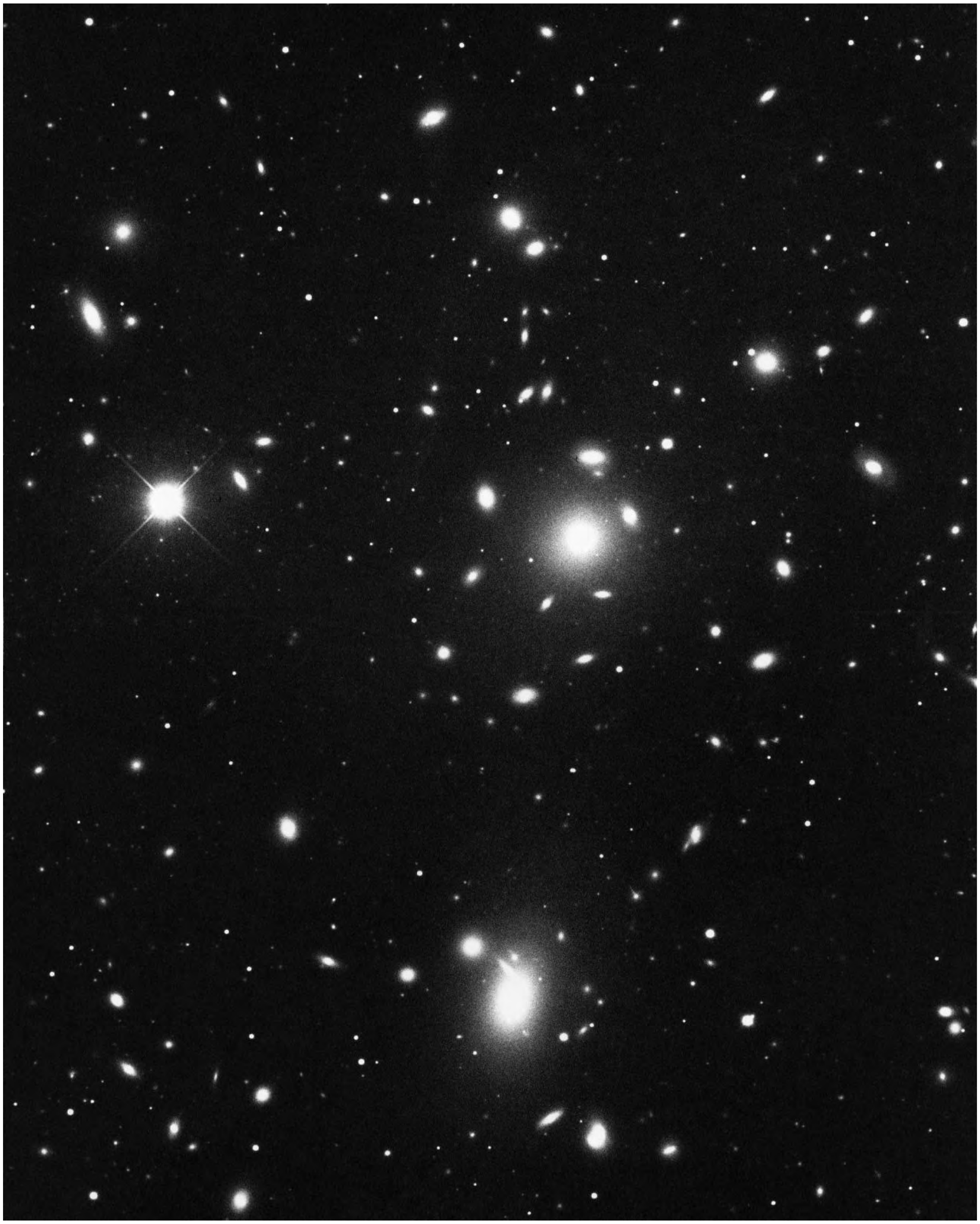


Figure 6.17: The Coma Cluster

A spherically shaped cluster of over 1000 galaxies, in the constellation of Coma Berenices, and located about 5 times further away than the Virgo Cluster. The Coma Cluster is dominated by elliptical and lenticular (S0) galaxies. The central giant elliptical is NGC 4889.

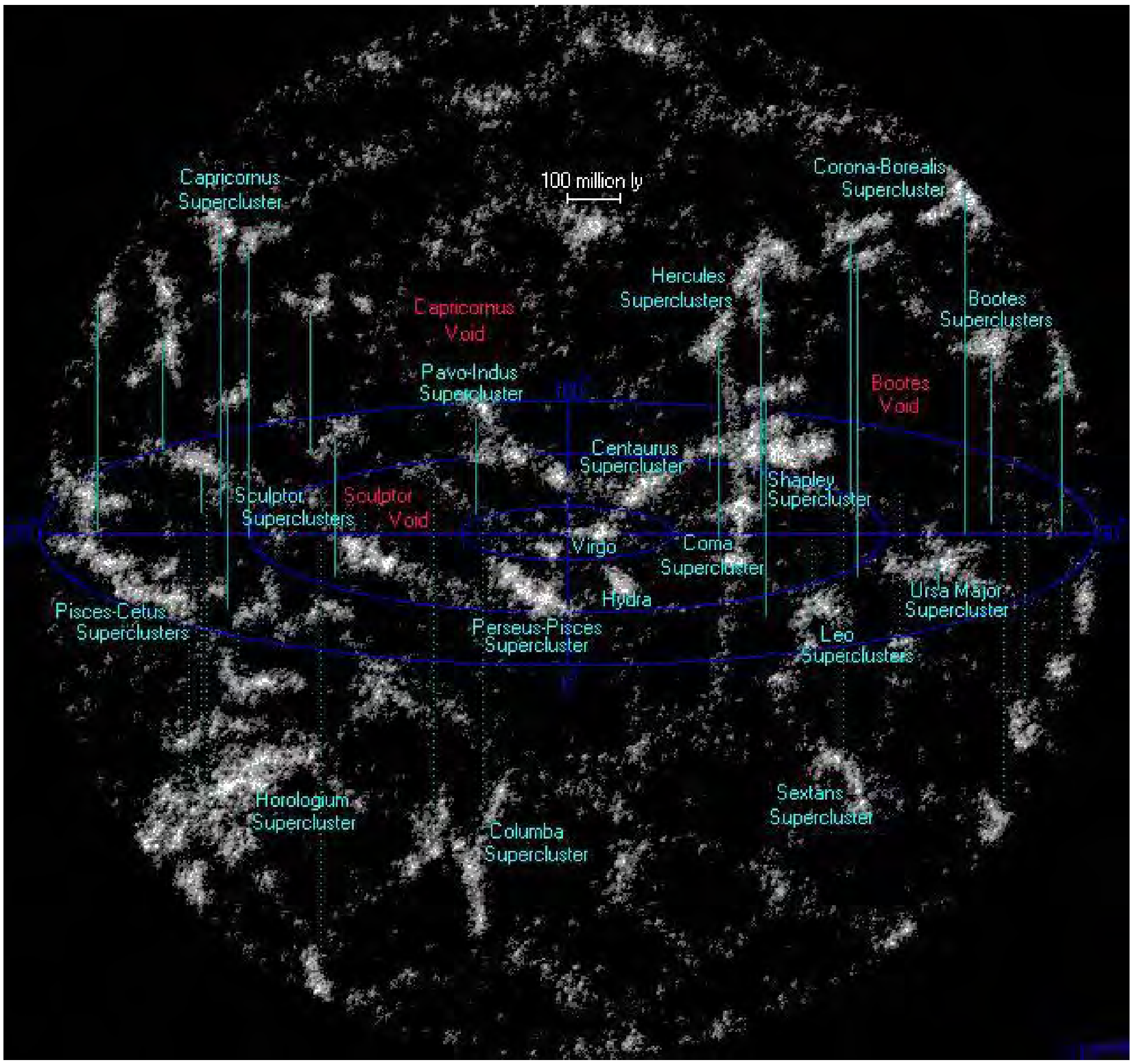


Figure 6.18: Neighboring Superclusters

Expanding out to a volume 1 billion light years in radius shows not only additional superclusters of galaxies, but we see now that galaxies at such scales are distributed unevenly in filaments and sheets, with vast voids in between.

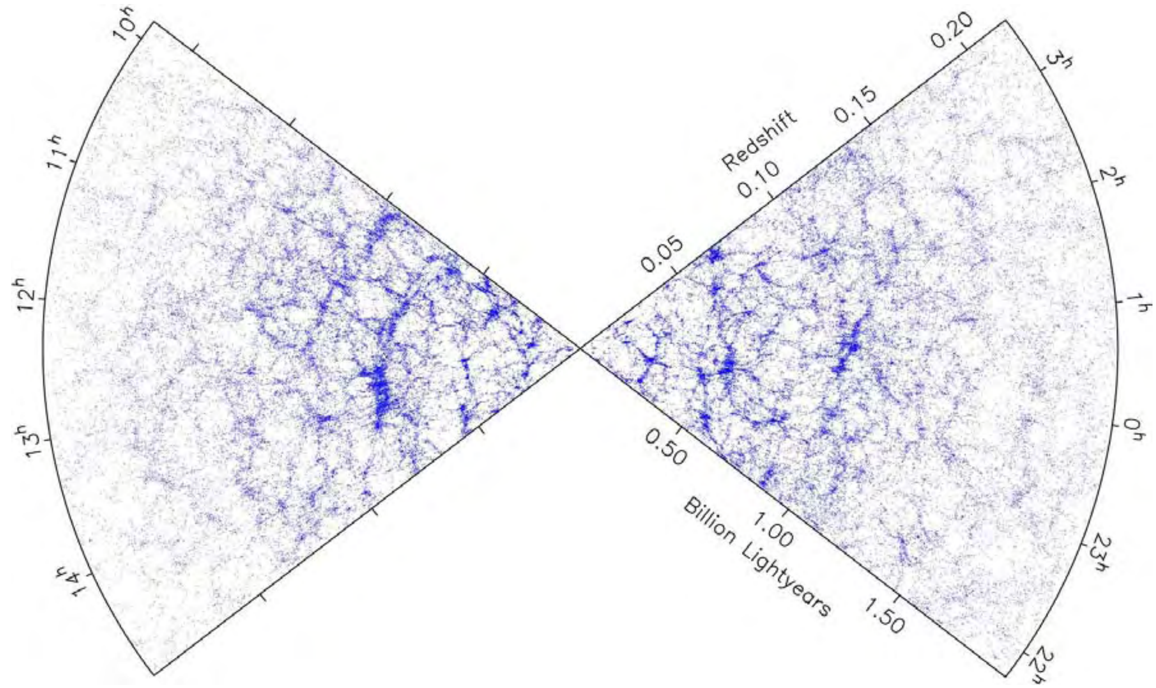


Figure 6.19: 2-Degree Field Galaxy Survey

Part of the data from the 2-Degree Field Redshift Galaxy Survey, containing positions for over 245,000 galaxies. In this image, the Milky Way is at the center of the two pie-shaped wedges. Astronomers observed two narrow strips of sky stretching out in the southern hemisphere (the right wedge) and the northern hemisphere (the left wedge). The galaxies are clearly distributed along giant filaments that surround enormous voids, like a mass of soap bubbles. The distribution of galaxies decrease the further we are from the center in this survey because the more distant galaxies are increasingly more difficult to observe.

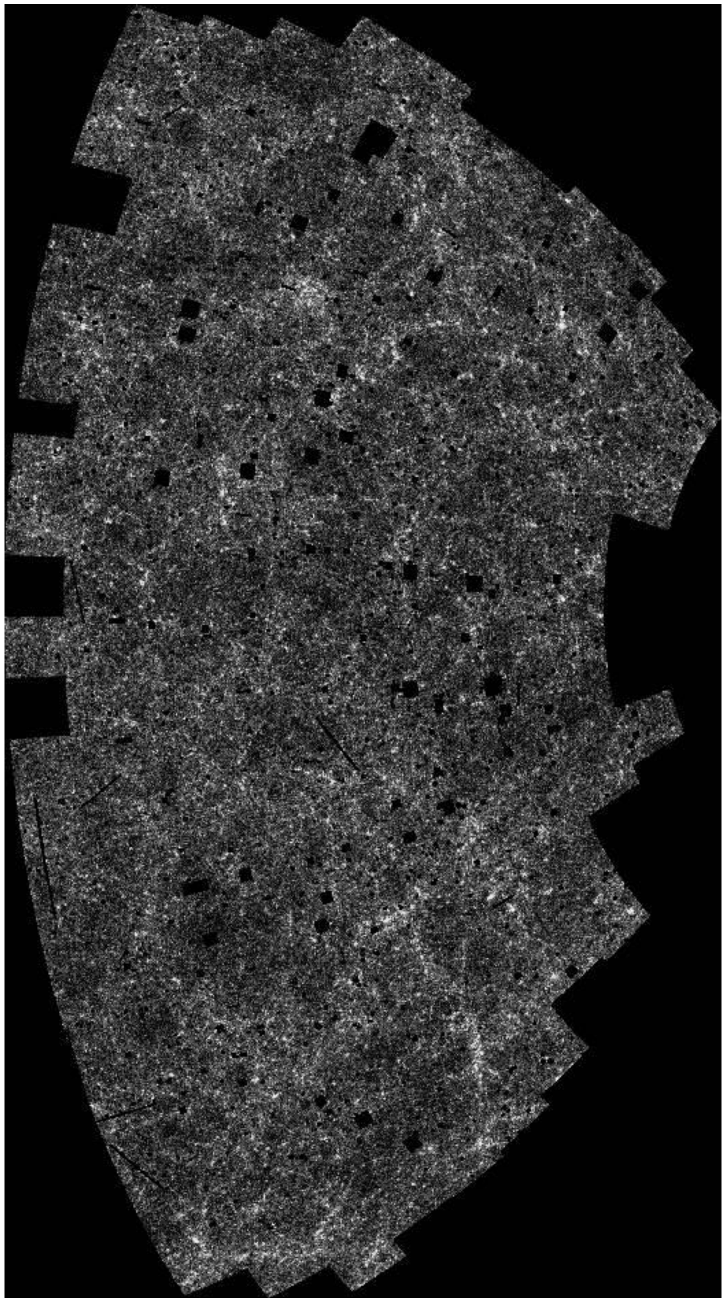


Figure 6.20: The APM Survey of Galaxies

Instead of trying to determine the 3-dimensional spatial positions of galaxies over a narrow strip of sky, this survey uses automated computer routines to look for galaxies in photographic plates, and determines their 2-dimensional location in the sky. About 4000 square degrees (1/10th of the entire sky!) are surveyed, with roughly 2 million galaxies found. The "holes" in the survey are regions containing bright stars that could not be scanned. The large-scale distribution of galaxies can be discerned in this map as well.

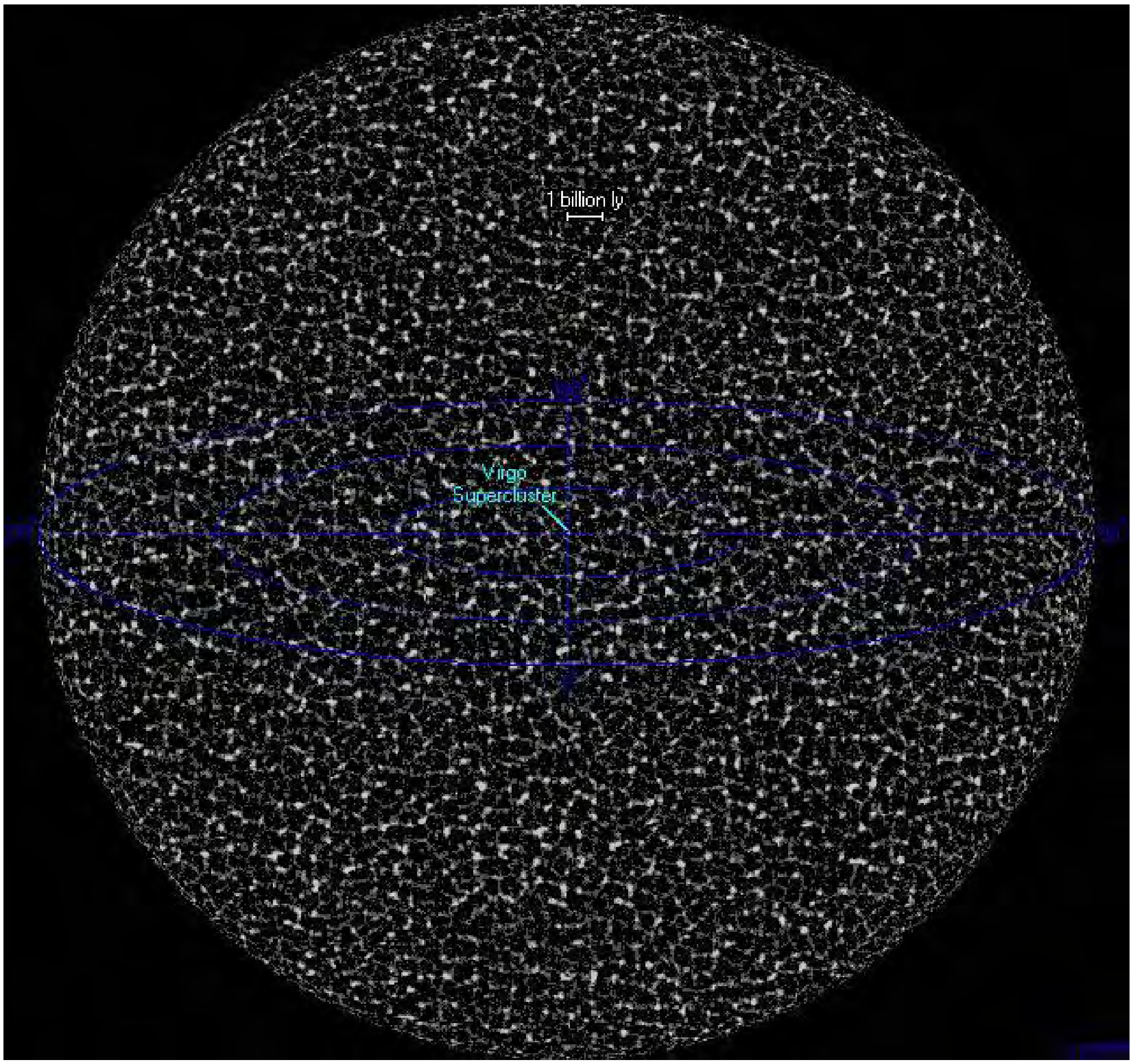


Figure 6.21: The Observable Universe
Making one more jump shows the edge of the observable Universe, roughly 14 billion light years in radius. The Universe, filled with vast voids 600 million light years across and superclusters of galaxies concentrated at the edges of these voids, finally starts to look uniform and homogeneous at such a scale.

ellipticals.

At larger scales, even larger **superclusters** of galaxies can be seen. The **Local Supercluster** is centered on the Virgo cluster, includes the Local Group, and is roughly 100 million light years across. The galaxy clusters that make up the supercluster are not gravitationally bound to each other, so the boundaries of a supercluster depend on the cutoff number density that separates regions of high galaxy count versus low galaxy count. At larger scales of 800 million light years, other superclusters are evident.

At the largest scales, superclusters appear to be organized in a vast network of sheets and filaments that surround nearly-empty, great voids. The **large-scale structure** of the Universe therefore appears to consist of chains of superclusters of galaxies, arranged around voids, of order 600 million light years across, like the pores inside sponges. The largest structures appear to have a scale of about 600 million light years. At even larger scales, the Universe finally starts to appear uniform or *homogeneous*—because any one 600 million light year sized region will look like any other 600 million light year patch.

6.1.4 Galaxy Cluster Mass

The mass of clusters of galaxies can be estimated in a number of ways. The first is to simply add up all of the light from the stars, gas, and dust in the galaxies. This gives the **luminous mass**, since this is mass emitting radiation that we can detect.

Another method is via the **virial theorem**. This theorem simply states that the distribution of velocities of a group of gravitationally bound objects will depend on their total collective mass. The greater the overall cluster mass, the faster the orbital speeds of the individual galaxies. When astronomers started applying the virial theorem to galaxy clusters, they found a very surprising result: the virial cluster mass was many times higher than the luminous mass. In fact, ordinary luminous matter in galaxies could account for only 20%–30% of the total cluster mass obtained from the virial method.

Another mass-finding method was to measure hot X-ray emitting **intracluster gas** that filled the space between galaxies. This is 10–100 million° K gas, which is extremely tenuous—several orders of magnitude less dense than the gas in the Solar corona. To determine the mass of the entire cluster, recall the concept of hydrostatic equilibrium for stars (see §5.4), which balances the weight of the star's gas envelope with the pressure from the energy generated in its core. A similar balance exists in galaxy clusters, where the pressure of the intracluster medium balances against the gravitational force of both the gas and galaxies. The measured temperatures and pressures of the intracluster medium again implies that the total gravitational field is *higher* than the gravity expected from the luminous matter.

Finally **gravitational lensing** has been used to derive the masses of galaxy clusters. Einstein's Theory of General Relativity predicts that the path that light takes

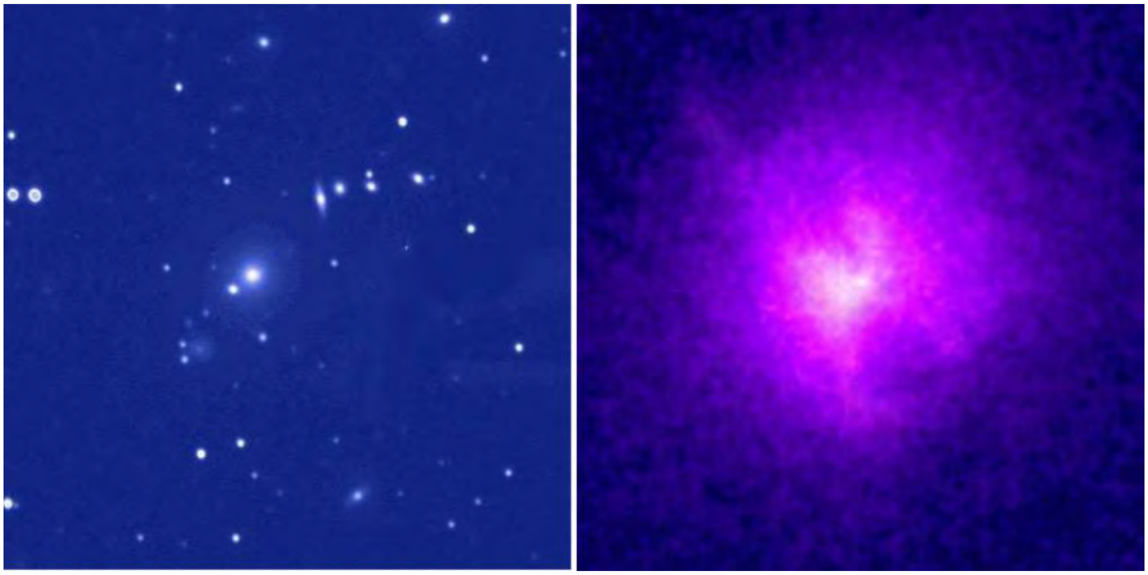


Figure 6.22: Hydra A Galaxy Cluster
A look at the Hydra A galaxy cluster in the optical (left) and in X-rays (right). The X-ray image shows an enormous bubble of extremely hot coronal gas that fills the inter-galactic medium in this cluster.

can be bent if it passes close by a massive object. Images of background galaxies are distorted because of the gravity of the foreground cluster, and the amount of the distortion can be used to infer the total cluster mass. Again the results are congruent with the virial and X-ray gas methods.

We come to the conclusion based on three completely independent techniques that *there is hidden matter that is gravitationally influencing the galaxies within the galactic clusters*. However this matter does not emit any radiation that we can detect! Depending on the estimate used, the total luminous mass of galaxies in a cluster is 10% or less, with the intracluster gas making up another 10–25%. The mysterious **dark matter** contributes 70–90% of the total cluster mass. Thus it appears that the majority of matter in the universe is not in any form that we can observe except by its gravitational effects!

6.1.5 Dark Matter

This problem of hidden dark matter was also apparent when spiral galaxies like our Milky Way were studied. One can measure the rotational speeds of stars in the disk of spiral galaxies and build up a **rotation curve** which plots the stellar velocities with their distance from the Galactic center. The velocities of stars within much of the disk are expected, given the observed luminous mass. However for stars located at the visible outer edge of the disk, the orbital velocity stays constant rather than

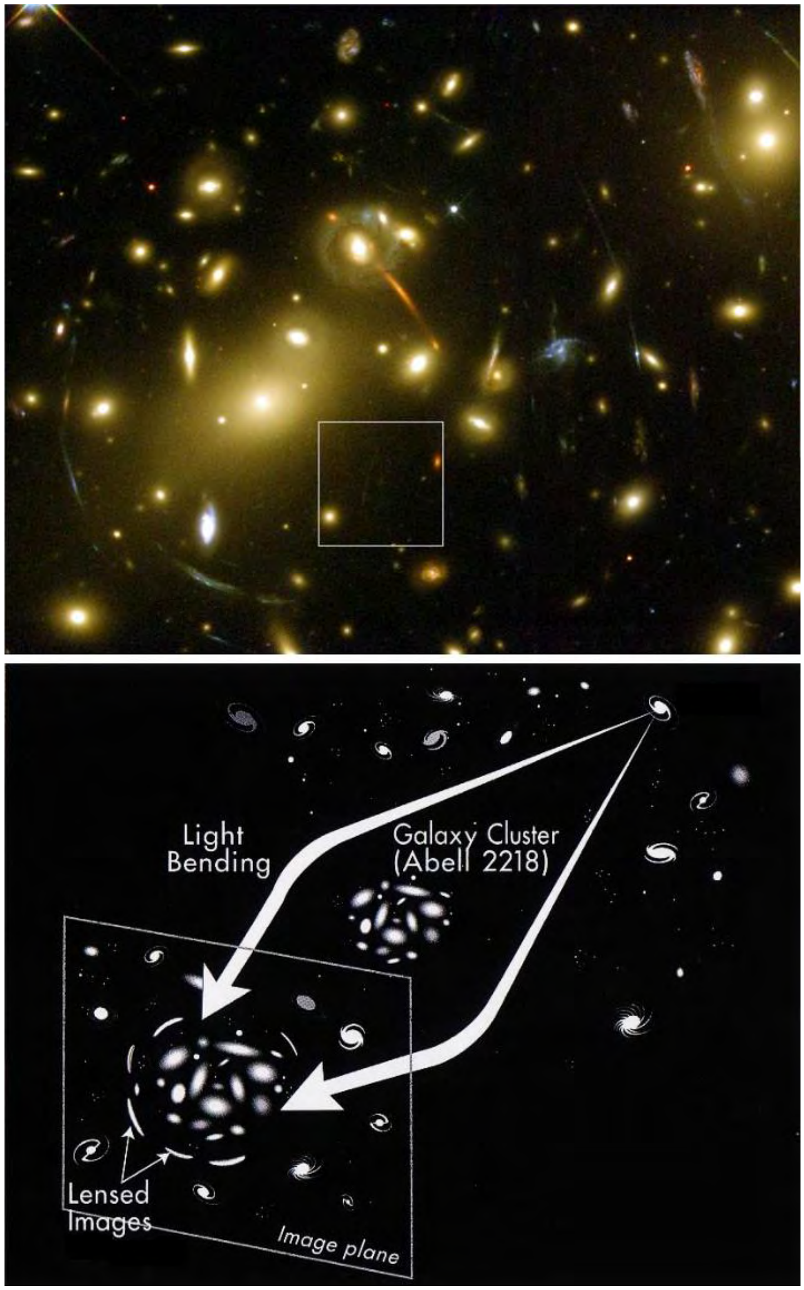


Figure 6.23: Gravitational Lensing by Abell 2218

The rich galaxy cluster Abell 2218, as viewed by the Hubble Space Telescope (top). The arcs that are centered around the bright elliptical just to the left of center are actually distorted images of a background galaxy. The diagram (bottom) shows how light from the background galaxy is bent by the gravity from the foreground cluster, arriving at our eyes (or our instrumentation) as circular arcs centered on the cluster.

falling off. This suggests that in addition to the visible matter in the outer reaches of a spiral galaxy, there is also a **dark matter halo**. Although the visible edge of the Milky Way's disk is located about 50,000 light years from its center, dark matter extends out at least to 65,000 light years. Depending on the assumptions used in setting the boundaries of the Galaxy and the distributions of the dark matter, the total mass of our Milky Way, including the dark matter, might be anywhere from 4–60 times the luminous mass from stars.

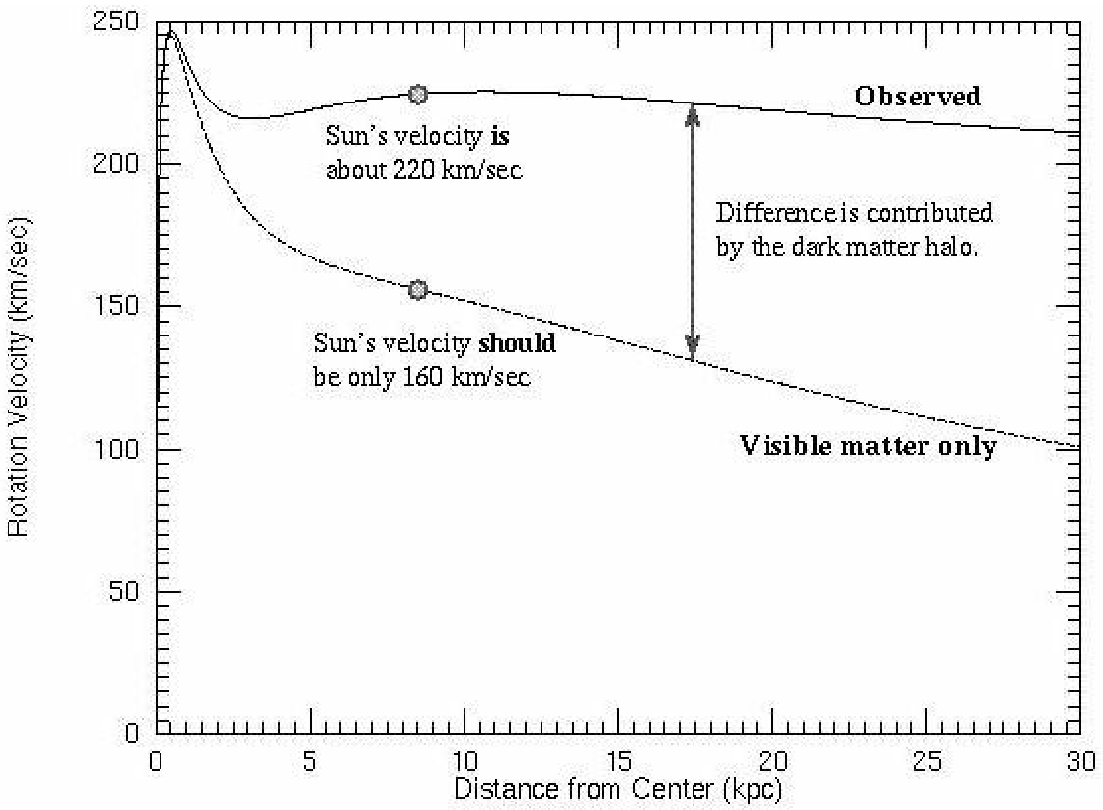


Figure 6.24: Rotation Curve of the Milky Way

A plot showing the orbital velocities of stars in the Milky Way (vertical axis) versus their distance (horizontal axis). Based on Newton's law of gravity, the more mass an object (like a star) orbits around, the faster its orbital speed. From a census of the visible matter in our Galaxy (stars plus gas and dust), we expect the velocities to follow the lower line. However the actual observed velocities of the stars (including our Sun) follows the top line, suggesting that there is more matter than can be observed. We therefore posit the existence of a dark matter halo that must be driving the orbital velocities up. The unit for distance in the plot is a kiloparsec (kpc), where 1 kpc is equivalent to 3260 light years.

What is the nature of dark matter? This is still under considerable debate, since we do not know enough about it. One conjecture suggests that the dark matter halo around the Milky Way could consist of ordinary matter in the form of **brown dwarfs**, star-like objects too small to start hydrogen fusion in their cores. These objects would be less than 80 times the mass of Jupiter, and would be too cool to radiate very much radiation—so little in fact that they would escape detection at both optical and infrared wavelengths. Another possibility is a population of stellar remnants, the end products of stellar evolution such as white dwarfs, neutron stars, and black holes. The white dwarfs and neutron stars would have stopped glowing after a few billion years. As long as no matter was available to be accreted, a black hole would not be visible. Attempts have been made to survey the halo for small, low mass objects via gravitational lensing, and such a population has been found. However the numbers of these objects are only about 20% of what is necessary to account for all of the dark matter halo mass.

The brown dwarfs, cold white dwarfs, and cold neutron stars are made up of ordinary matter that we are familiar with (although white dwarfs and neutron stars are far denser than matter that we are used to). The majority of their mass is in the form of neutrons and protons, which are collectively called **baryons**. (The electrons are 1/1000th as massive by comparison, so they do not contribute appreciably to the mass budget.)

In addition to this relatively cold, baryonic dark matter, astronomers have also suggested that **non-baryonic** matter could make up the bulk of the dark matter. This is the so-called **hot** dark matter component, because such particles would move at velocities very close to the speed of light. The neutrino has been mentioned as a candidate for such a particle. Although neutrinos were thought to have **zero rest mass**, and would therefore travel at the speed of light, physicists realized that there was no fundamental reason why they should be massless. Therefore if neutrinos did have a very small mass, they could provide some fraction of the dark matter mass by sheer numbers. Recent experiments involving neutrinos from the Sun showed that they have a mass about 5 million times less massive than the electron. This however is not enough; the neutrinos would constitute on order about 1% of all dark matter.

Physicists have also suggested **WIMPs**, the **weakly interacting massive particles** as a dark matter candidate. The "weak" in the name refers to one of the four fundamental forces in the Universe that affects such particles.[6] According to proposed **supersymmetry** theories of particle physics, relationships between fundamental particles and the forces of nature implies the existence of new classes of undiscovered particles which interact only by gravity and the weak force. One such proposed particle, the **neutralino**, has a mass 20–1000 times that of a proton. It is expected that if supersymmetric theories are correct, many neutralinos would have been created in the early universe. Currently there are efforts (unsuccessful so far) to generate artificially neutralinos in particle accelerators, as well as to detect any cos-

[6]The other three forces are: the **strong force** which mediates interactions between protons and neutrons in atomic nuclei, and keeps them bound together; the **electromagnetic force** which controls interactions between charged particles; and **gravity**.

mic neutralinos that might be wandering through the Solar System from the Galactic halo.

6.2 The Big Bang

6.2.1 The Expanding Universe

There are a handful of observations that give clues to the origins of the Universe. First is the **cosmological expansion**: *the entire Universe is expanding so that matter is spreading apart.* Some structures, like galaxy clusters, are gravitationally bound together, and resist the expansion. But at size scales of superclusters (hundreds of millions of light years) or larger, the galaxy clusters tend to be weakly bound and are following the general expansion of space, and are flying further apart from one another.

The chief observational evidence for this expansion is the **redshift** of galaxies. As part of its wave-like properties, light can be **Doppler-shifted** when the source of the light is moving toward or away from an observer. An familiar, everyday example is the siren of an ambulance or fire engine which is pitched *higher* when the vehicle is moving toward you, and pitched *lower* when it is moving away. The higher pitch is the result of the signal being **blue-shifted**, or shifted to higher frequencies (and shorter wavelengths). A source moving away will have its signal shifted to lower frequencies, or **red-shifted**. Similarly when observing galaxies, nearly all of them are moving away from the Milky Way. And just as significant, the *distances to the galaxies are proportional to their red-shifted velocities.* Thus the further away the galaxy is, the greater its speed of motion away from us.

This bizarre behavior of the galaxies might suggest that our Milky Way is special, that every galaxy in the Universe is fleeing from us. However as we have seen in §6.1.3, at size scales on the order of 600 million light years, the universe is **homogeneous** and **isotropic**—looking the same in every direction. Any part of the Universe is more or less like any other part. Recall the Copernican principle that we do not hold a privileged place via the Earth's place in the Solar System, nor the Sun's place in the Milky Way. It seems natural to continue extending it one step further to include our Galaxy's place in the entire Universe. Therefore a more likely explanation for the recession of the galaxies around us is that the *entire Universe is expanding*, and regardless of which galaxy you reside in, it seems as if every other galaxy is moving away from you with a speed proportional to its distance.

Since the Universe appears to be expanding, what happens if we run the movie of the cosmos backward? The Universe would *contract.* If we keep rewinding the film, the density and temperature of the matter will continue to rise, until they reach extraordinarily high values as the contraction continues toward a single point. This point is the explosive origin of not only all matter and energy, but of space-time itself in the Universe, and hence is referred to as the **Big Bang**.

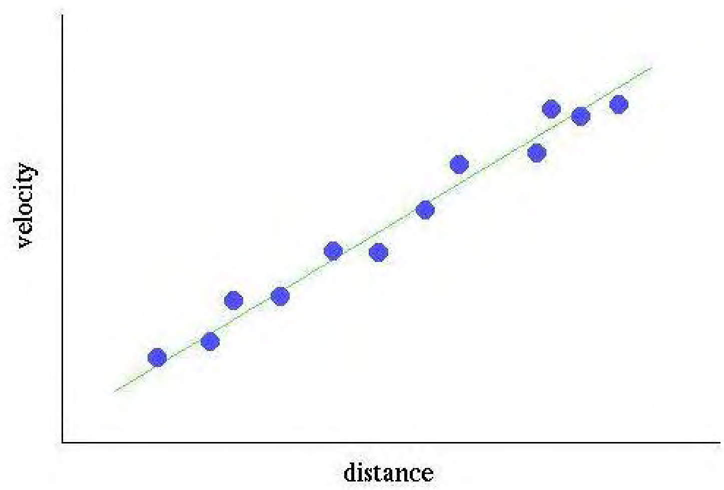


Figure 6.25: Hubble's Law
Galaxies are expanding away from us. As this plot of a sample of galaxy velocities show, the further they are in distance from us, the greater their expansion velocity. The velocity-distance relationship is not exact, since galaxies are found in clusters and the gravitational interactions between members of clusters can alter a galaxy's velocity from the mean Hubble flow of expansion.

6.2.2 Cosmic Element Abundances

In the first few minutes after the Big Bang, the temperature and pressure conditions were such that there were only free protons, neutrons, electrons, along with electromagnetic radiation in the form of photons. As the temperature dropped below 100 million° K, **primordial nucleosynthesis** became possible: the creation of atomic nuclei in the aftermath of the Big Bang. The important difference between this and the *stellar nucleosynthesis* that is still occurring today inside stars (as discussed in Lectures 4 and 5) is that the conditions after the Big Bang were constantly changing. As the Universe grew larger after the Big Bang, the pressures and temperatures dropped. The intense initial temperatures meant however the nuclear reactions occurred at rates much higher than that found inside stars.

The first important nuclear reaction created **deuterium** or ^{2}H:

$$p + n \rightleftharpoons {}^2\mathrm{H} + \gamma,$$

where γ represents a gamma-ray photon. This reaction is *reversible* (and hence the arrows pointing left and right) meaning a sufficiently energetic photon can destroy the deuterium nucleus. However once the temperature of the Universe dropped below 1 billion° K, the average photon energies dropped to the point where quantities of deuterium could build up. There are no known processes in the present Universe that can create substantial amounts of deuterium; stars can only destroy deuterium as a fuel source in reactions at their cores. *All of the deuterium today is the result of these primordial reactions.* As a result, there are many studies underway today to determine the exact amount of present day deuterium in order to understand the conditions of the early Universe.

After deuterium, the early Universe cooled enough to start producing Helium-4 (^{4}He) in large numbers. A small amount of Lithium-7 (^{7}Li) was also created before the temperature dropped too much. After about 1000 seconds (17 minutes) after the Big Bang, the temperature was below 500 million° K and all element creation ended.

The computer simulations of this early period in our Universe showed that the total amount of ^{4}He created should be about 24%. This agrees remarkably well with observations of not just interstellar gas not polluted by heavy element enrichment from stars, but with the rough proportion of observed helium abundance *everywhere* in the Universe. That the Big Bang (and not any other model) can predict elemental abundances so close to what is actually observed is one of its underlying strengths as a theory.[7]

6.2.3 The Cosmic Microwave Background

During the next 300,000–400,000 years, the temperature continued to drop in the Universe as it expanded. The composition of the Universe consisted of photons, neutrinos, protons (which would become hydrogen nuclei), helium nuclei, a smattering of deuterium and lithium nuclei, electrons, and dark matter particles (whatever they might be!). The photons collided with and scattered between the free electrons, sharing energy. The electrons also collided and shared energy with the baryons, the protons and atomic nuclei. All of the protons, nuclei, electrons, and photons could therefore be described as having the same temperature.

However once conditions cooled to 4500° K, the era of **recombination** started. Electrons started combining with atomic nuclei and free protons to form neutral atoms. By the time the temperature had dropped to 3000° K, most of the electrons were bound up in atoms. There were not enough free electrons for the photons to scatter off of, and the Universe became *transparent*. The photons could now stream freely, instead of interacting with the electrons and nuclei. Because the photons could no longer easily share their energy with other particles, they became **decoupled** from the matter. The matter and photons now evolved separately from each other.

One consequence of the cosmological expansion of the Universe was to cause the wavelengths of the photons to get stretched out. The photons therefore became **red-shifted**. Because the photons were originally coupled to the dense hot matter soup (containing protons, electrons, and heavier nuclei), they retained a blackbody spectrum. After decoupling, the blackbody spectrum continued to red-shift to lower temperatures and longer wavelengths. Today this radiation makes up the **cosmic**

[7]See Appendix B to see what scientists mean when they call something a “theory”.

microwave background, with a temperature just a few degrees above absolute zero (which is why it emits mostly in the microwaves).

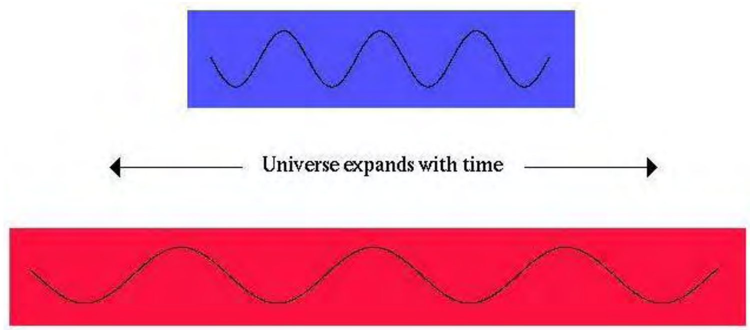


Figure 6.26: The Expanding Wavelength of a Photon
The wavelength of a photon traveling through space expands with the Universe. Because the wavelength gets longer, the photon will get redder *over time. This explains why photons that originated from the last-scattering surface when the Universe was still extremely hot are now affiliated with a blackbody temperature of less than 3° K.*

The cosmic background radiation consists of photons coming from the **last-scattering surface**. These photons last encountered another matter particle when the Universe was opaque, just before recombination. Immediately after recombination, they were able to free-stream through space to reach our detectors. Although this last-scattering surface appears as a spherical shell (like the celestial sphere) that surrounds us, this *does not* mean we are at the center of the Universe. A useful analogy is if you go out walking on a foggy day and there is a visibility of 50 feet: light from objects does not reach you from beyond a distance of 50 feet. Instead the photons are scattered by the suspended water droplets that make up the fog. As you walk around this "foggy universe," it will feel as if the universe has a radius of 50 feet, defined by how far you can see. Another observer taking a separate walk outside will see a different last-scattering surface. Each observer will have her own *observable universe* that is slightly different than any other observer's. *Thus there is a limit to how far we can observe in the Universe!* The Universe might be much larger than our **visible horizon** as defined by the cosmic microwave background, but there is no way for us to detect it.

The cosmic microwave background is the third observational triumph that supports the Big Bang theory. It was predicted by theoretical calculations in the 1940s, and was first detected as a uniform, isotropic microwave radio emission in 1965 by Arno Penzias and Robert Wilson of Bell Laboratories. Throughout the 1970s and 1980s, ground- and balloon-based measurements were able to show that the cosmic microwave background had a blackbody-like spectrum. In 1989, the Cosmic Microwave Background Explorer (COBE) was launched into space and it was able to

Figure 6.27: Seeing to the Edge of the Universe

Although we are looking backward in time when we peer out into the deepest reaches of the Universe (since light has a finite velocity), this does not *mean that we can see all the way to the Big Bang. Photons before the time of recombination were scattered like light in a deep fog. We can only detect photons that were created after the period of recombination (or light coming from objects closer than the fog limit.*

make detailed, unprecedented measurements of the characteristics of this emission. Not only did it confirm the uniform and isotropic nature of the emission, but it found the blackbody spectrum to have a temperature of $2.725 \pm 0.002°$ K. The blackbody spectrum that it measured spectroscopically was perhaps the most perfect example of a blackbody ever seen. No other emission in nature or from laboratories on Earth have spectra that matches so closely to that of a theoretical blackbody spectrum.

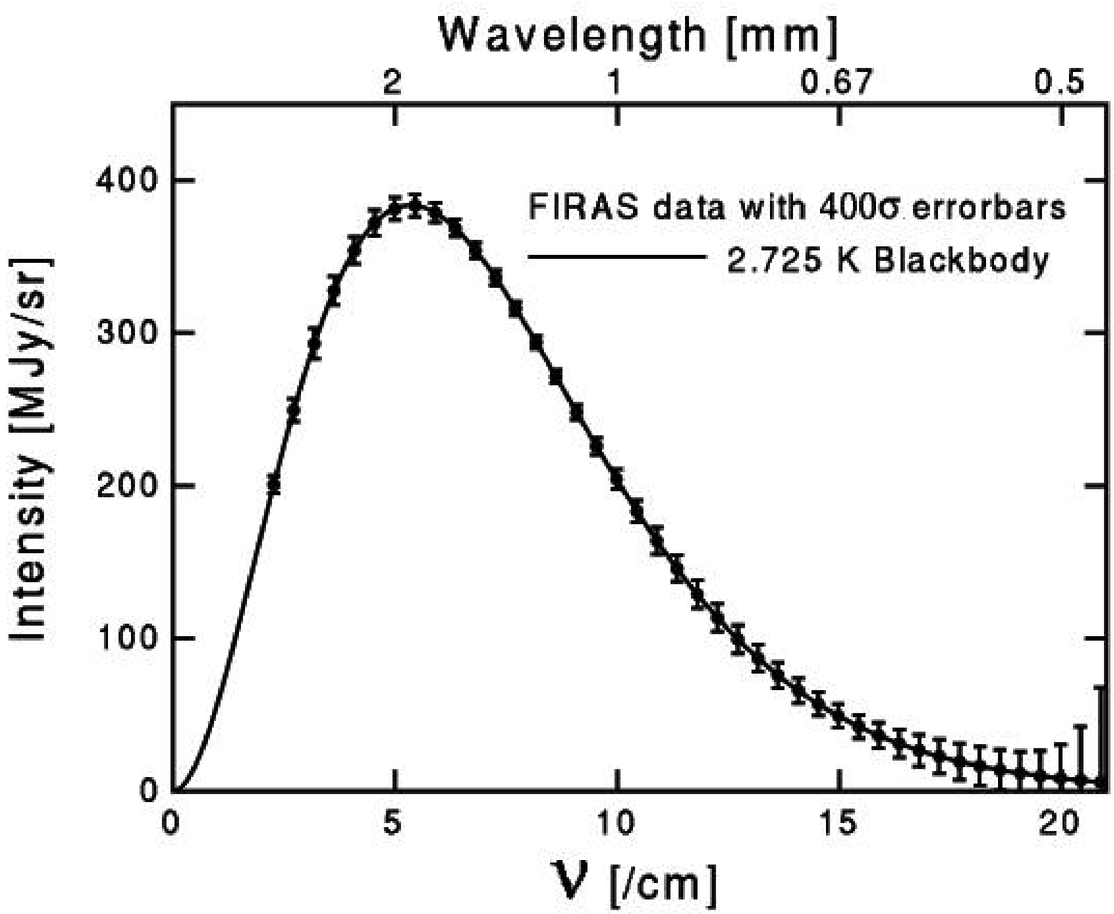


Figure 6.28: COBE Spectrum

The spectrum of the cosmic microwave background as measured by the COBE spacecraft. The dots are the actual measurements. The error bars of the measurements are shown 400 times larger than their actual size*; normally they would be undetectable in this image! Finally the smooth curve is a theoretical blackbody spectrum at a temperature of 2.725° K, which the observations show a nearly perfect agreement with.*

6.3 Inflation

6.3.1 Problems With the Big Bang

Despite the successes of the Big Bang theory, there were still some problems that were unresolved in the basic model. Recall that in the last section, the cosmic microwave background was found to be extremely uniform over the entire sky. In fact it is uniform to a few parts in 10^5, meaning there is less than a 0.0001° K difference in temperature from one part of the sky to another. However at the time of last-scattering, when the photons were coupled to the matter, the closest neighboring photons that could exchange energy with each other at the speed of light is equivalent to only 2° on the sky today. This is the **horizon problem**: how can parts of the cosmic microwave background 180° apart exchange energy to have the same temperature today? Somehow these very widely separated parts of the Universe today must have been far closer together to be able to swap energy and reach thermal equilibrium, then is suggested by the normal Big Bang model.

The next issue is the **flatness problem**. There are three possible outcomes to an expanding universe, and it depends on what the density of the Universe is. The Universe might be over-dense, meaning it has enough matter in it to stop the expansion, and then cause the Universe to contract toward a Big Crunch. This is a "closed" universe. The Universe might have far less mass in it to stop or even slow down the expansion. In this case the Universe would continue to expand faster and faster. This is an "open" universe. Finally the third scenario is the critical case in between where there is just enough matter to stop the expansion, but only after an infinite amount of time. The Universe will therefore expand more slowly over time, but will never completely stop. This last case is known as a "flat" Universe, or the $\Omega = 1$ scenario, where Ω is the critical ratio between the matter in the Universe and the amount of matter necessary to stop the expansion. (A Universe where a Big Crunch will occur is $\Omega > 1$, while a Universe that will expand faster over time is $\Omega < 1$.) The terms *closed*, *flat*, and *open* refer to the geometry of space-time in the Universe. This four-dimensional depiction of the Universe is a consequence of Einstein's Theory of General Relativity.

By almost all measurements, it seemed that Ω was very close to 1, if not in fact equal to 1. However this presents a problem: if Ω is nearly 1 today, it would have to be even closer to 1 in the past. This is because if Ω was slightly different than 1 in the past, this difference would have grown over time. According to one model of the expansion, if $\Omega = 0.90$ today, then when the Universe was 1/30th its present age, $\Omega = 0.99$. However when the Universe was only 1 second old, Ω must have been even closer to 1: 1.000000000001. At even earlier times, Ω would have been even *closer still* to 1. How did the Universe come to be so perfectly tuned? One might say it just happened this way, but that is not a very satisfactory answer to cosmologists.

According to certain theories of particle physics, the conditions in the Big Bang

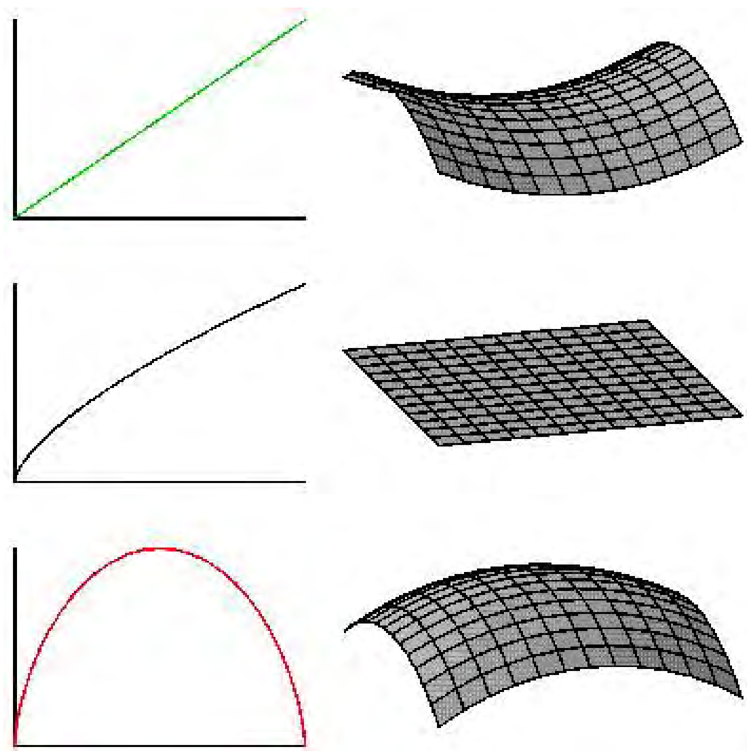


Figure 6.29: Curvature of Three Possible Universes
Given that it is already expanding, the Universe has three possible end scenarios. The left plots show how the size of the Universe changes over time. (Top Left) *The expansion continues forever at a constant rate.* (Bottom Left) *The expansion slows to a stop, before reversing course and the Universe re-collapses back on itself.* (Middle Left) *An intermediate state also exists where the expansion continues forever but the expansion slows down over time, without ever stopping. Which of these describe the Universe depends on the total amount of matter in it, and whether the collective gravitational pull of this matter can slow or even halt the expansion. The pictures on the right show the equivalent "geometry" of space-time in the three scenarios based on the total matter (from top to bottom): an open or negatively curved geometry, a flat geometry, and a closed or positively curved geometry.*

were expected to create a slew of unusual objects, including **magnetic monopoles**, gravitinos, and other exotic subatomic particles. These are massive particles that do not decay and should be easily found in the present day Universe. "Topological" defects in space-time itself, such as "cosmic strings" and "domain walls" are also expected. In fact the mass of the magnetic monopoles alone would be enough to close the observable Universe. However no magnetic monopoles have ever been detected.

One last problem with our model of the Universe is that there appears to be an imbalance between matter and antimatter. Matter will annihilate perfectly with its antimatter counterpart (electrons with positrons, protons with anti-protons, neutrons with anti-neutrons, etc.) with high energy photons as a result. Similarly the reverse reaction can occur: photons can spontaneously create matter-antimatter pairs, such as an electron and a positron. One would there expect the early Big Bang to produce equal amounts of matter and antimatter in the Universe. However by all accounts, there is very little antimatter. The Universe is mostly matter.

6.3.2 Problems Solved?

One answer that would solve all of these problems is **inflation**, first proposed by physicist Alan Guth in 1980. This hypothesis suggests that the universe went through a period very early in its history where it expanded at an exponential rate, about $t = 10^{-36}$ second after the Big Bang. The origins of this theory came from an analysis of the properties of the vacuum. The early Universe may have had a **false vacuum** with an energy density inherent in the vacuum much higher than the **true vacuum** that exists today. As the Universe cooled to a temperature of $10^{28\circ}$ K, the vacuum changed from one form to another. But before this **phase change** could finish, the Universe expanded at a much, much higher rate than either before or after this transition. Guth showed that between $t = 10^{-36}$ second and $t = 10^{-34}$ second after the Big Bang, his inflation theory predicted the size of the Universe to increase by a factor of 10^{50}!

This enormous exponential expansion solved many of the problems we mentioned in the last section. The horizon problem is no longer an issue: The regions now separated by great distances in the present epoch were actually once much closer together. This gives a chance for the radiation to come to the same temperature everywhere in the observable Universe, before it was spread apart by inflation.

Inflation solves the flatness problem by expanding the spacetime geometry of the Universe. Even if the Universe had not started out being flat, the vast rate of expansion stretched the curvature of Universe until it appeared flat. A similar analogy is that a small balloon might have a curved geometry. However if you were to blow up the balloon until it was the size of the Earth, anyone standing on the surface would see a flat surface. (If inflation were to occur for our balloon, it would blow up to a size about 10^{24} times larger than the current size of the observable Universe; the surface of such a balloon would indeed look extremely flat!)

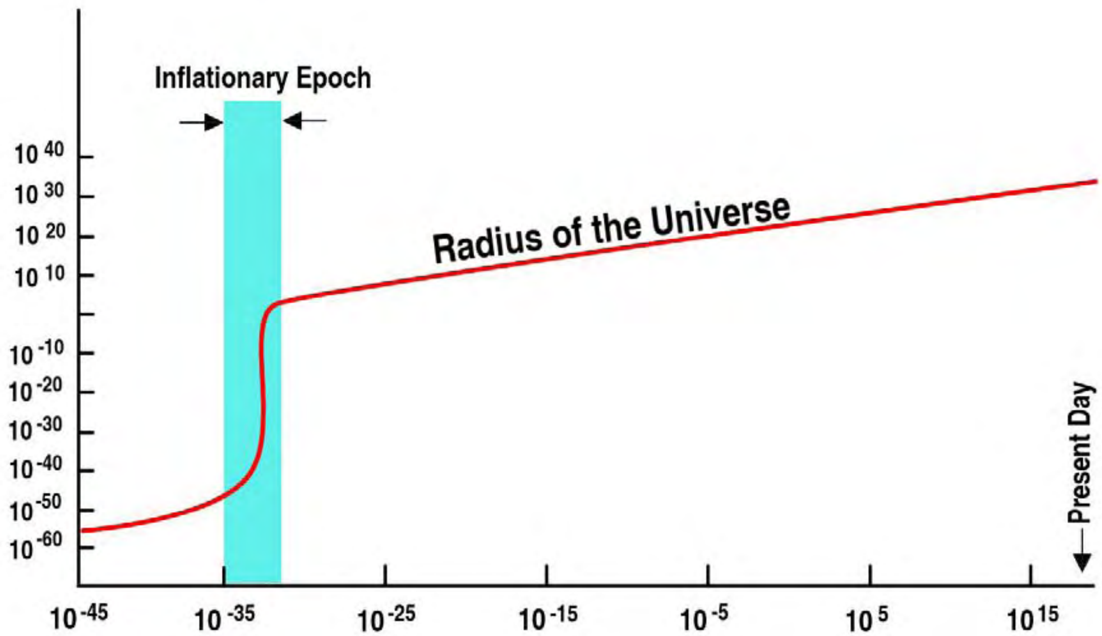


Figure 6.30: Expansion from Inflation

The inflation hypothesis supposes that the Universe underwent an enormous expansion about 10^{-35} *seconds after the Big Bang. According to some versions of the model, it might have gotten larger by a factor of* 10^{50} *times (!). After lasting for less than* 10^{-30} *seconds, the inflationary period (cyan) ended, and the Universe continued to expand at a much more leisurely rate up through the present day.*

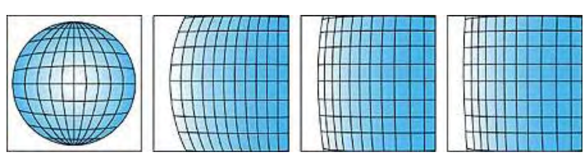


Figure 6.31: Inflation Solves the Flatness Problem

Even if the initial Universe had an extremely curved space-time geometry, the inflationary period blew up the Universe to such a great extent that today, it appears flat to us. Similarly if a sphere the size of a proton were to expand until it was the size of the Earth, its surface to us would look extremely flat.

The problem of magnetic monopoles and other exotic objects is also solved by such an enormous expansion. The original pool of magnetic monopoles would have expanded along with the rest of the Universe. The expansion is so great that they would be so spread out, that there would likely not be a single magnetic monopole left within our observable Universe.

Inflationary theory does not quite solve the imbalance problem between matter and antimatter. However another theory involving the **unification** of the strong nuclear force with the electromagnetic and weak forces suggests that at high enough energies, matter could have formed with a slight excess over antimatter. As the Universe cooled after this very early period to the time $t = 10^{-5}$ second, all of the protons annihilated with the anti-protons to create gamma ray photons. For every billion (10^9) proton-anti-proton pairs, there was an extra proton that *did not* annihilate. These residual protons would eventually become all the matter that we see today. The photons that came out of the annihilations became the radiation bath that expanded with the Universe and turned into the radiation background discussed in §6.2.3.

Despite all these apparent successes, inflation is in no way as robust a model as the Big Bang theory. In fact, there is not one inflationary hypothesis but a class of many competing theories. Alan Guth's original model is not even under consideration any more because of problems that were discovered with it. Even the exact mechanism for initiating inflation is not well understood. Although inflation solves many fundamental problems of the Big Bang, this feature does not guarantee that inflation actually occurred. The very ad hoc nature of the many variants of inflation—people often came up with new variations to explain different observations—may even argue against the idea as a whole. However many physicists would argue that the *concept* of inflation has solved far more problems than it introduces, and some version of it is here to stay.

6.4 The Contents and Fate of the Universe

There are a number of fundamental parameters that effectively describe our Universe. One is the **Hubble constant**, which measures how fast the Universe is expanding. Great effort has been expended ever since Edwin Hubble's time in trying to determine the expansion rate, and recent efforts have managed to narrow the error bars associated with the various measurements. This Hubble constant is however the *current* rate of expansion; the expansion might be slower or faster in the past. Determining either the acceleration or deceleration of the expansion over time will enable us to learn the eventual fate of the Universe. If the expansion slows but does not completely stop, the Universe will effectively expand forever. If the expansion slows eventually to a stop, then the Universe will contract into a Big Crunch. One can make arguments that the $\Omega = 1$ (or at least very close to it according to §6.3) which would mean a perpetually slowing expansion. What do the observations tell us?

One attempt to measure the deceleration parameter uses observations of supernova explosions. The supernovae that we learned about in Lecture 5—which are the result of the implosion of a single massive star—are **Type II** supernovae. Another type of supernova can occur in a *binary star* system with two stars of roughly solar mass. One star expands into a red giant and turns into a white dwarf. The second star is slightly less massive so it will expand into a red giant later. Once it does, the stellar envelope of the second star accretes onto the white dwarf. As noted in §5.6, the upper limit for white dwarf mass is 1.4 $M_\odot$. Any object that is larger will collapse into a neutron star with a supernova explosion as a result. The accreting white dwarf will therefore collapse into a neutron star exactly at the moment that it reaches the critical mass of 1.4 $M_\odot$. Such a **Type Ia** supernova will explode with the same amount of energy for every binary white dwarf-red giant system, and will have the same maximum brightness. They can therefore be used as **standard candles**: if two Type Ia supernovae have different brightnesses, *the fainter one must be further away than the bright one.*

Two teams of astronomers worked on using the Type Ia supernova method to search for deceleration. What they found starting in 1998 was surprising. The most distant supernovae appeared to be *closer* than what a steady, constant expansion of the Universe would have suggested. This suggests that the expansion has increased, not decreased over time. *The Universe is accelerating in its expansion!*

The possibility of a force speeding up the expansion of the Universe (or acting in reverse) has actually been recognized for more than eight decades. After Einstein came up with his Theory of General Relativity, he realized that his equations could be used to describe the history of the Universe. However his model Universe was not static: it would either collapse or expand over time. The expansion of the Universe had not been discovered yet, so Einstein added a **cosmological constant**, signified by Λ (the Greek letter 'lambda'), into his equations to keep the Universe static. Note that a cosmological constant can be either negative or positive, which would either prevent expansion or prevent a collapse. However once evidence for an expanding Universe was found by Hubble, Einstein (as well as subsequent physicists and astronomers) had no need for the constant anymore. Up until the supernovae studies, it was assumed not only that $\Lambda = 0$, but that $\Omega = 1$ (by the arguments from §6.3), and the contribution to the value of Ω was from matter, either ordinary baryonic matter or more exotic dark matter. We may therefore write:

$$\Omega = \Omega_m = 1,$$

where Ω_m is the amount of matter that is necessary to critically keep the Universe flat. One problem with this proposal was that based on all observations, there was not enough of *any* kind of matter to keep the Universe flat. The total amount of ordinary matter could be tallied up from observations, as shown in Table 6.1.

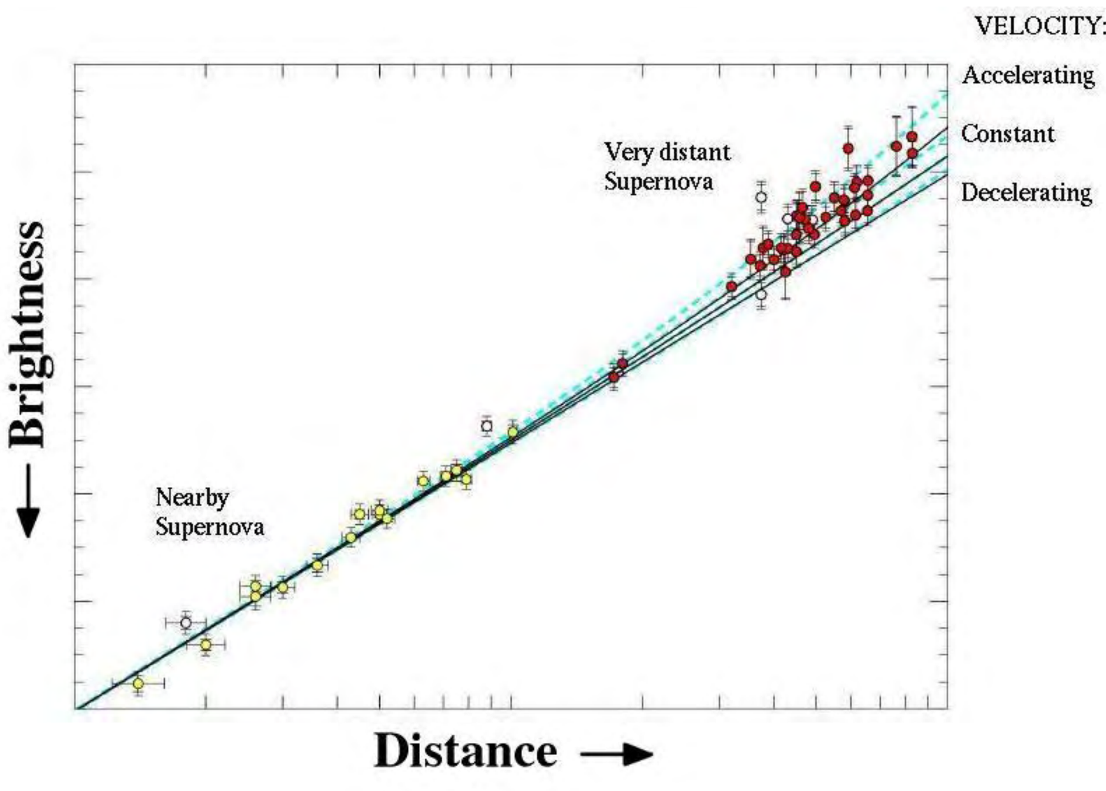


Figure 6.32: The Accelerating Universe

By measuring the brightness of distant Type Ia supernovae, teams of astronomers have found that they appear to be further away than where they should given the current expansion of the Universe. One conclusion that can be made from this evidence is that we live in a Universe that is not only expanding, but the expansion is accelerating as well.

Table 6.1: Baryonic Census

Source	Contribution to Ω
elliptical galaxy stars	0.0026
spiral galaxy stars	0.0009
irregular galaxy stars	0.0001
neutral atomic gas	0.0003
molecular gas	0.0003
ionized gas in galaxy clusters	0.0082
ionized gas in cool clouds	0.0020

Depending on the error bars associated with the various measurements, the contribution to Ω from baryons would be in the range

$$0.007 \leq \Omega_b \leq 0.041$$

with a "best guess" of:

$$\Omega_b \approx 0.021.$$

As for the determination of the total mass from matter (dark as well as baryonic), recall in §6.1.4 the various means used for determining the total gravitating mass in galaxy clusters. Using a combination of results from these methodologies, it seems the total matter contribution to Ω was only

$$\Omega_m = 0.16$$

However the results of the recent supernovae projects suggest that a cosmological constant might indeed be real! In fact a cosmological constant can add to the total energy density of the University. The Type Ia supernovae results show that:

$$\Omega_\Lambda \approx 0.70$$

which implies a slightly different Ω_m:[8]

$$\Omega_m \approx 0.30.$$

These results for Λ have been supported by subsequent (and completely independent) work involving the study of the fine details of fluctuations in the cosmic microwave background. It has become increasingly clear then that in addition to dark matter, there is an even more mysterious **dark energy** that pervades the cosmos and causes an acceleration of its expansion.

[8]This is reasonable given the considerable amount of uncertainty that exists for many astrophysical observations. A factor of two difference from Ω_m is actually not too bad!

Physicists are already working on models that could explain the workings and origins of dark energy. It is still so early that none are very plausible or satisfying at the moment. However the discovery of dark energy makes a fascinating denouement to our study of astronomy and our place in the cosmos. From the time of the Renaissance and the beginnings of modern science, the Copernican principle has lifted its head periodically to put humanity in its place. Its first manifestation occurred when observations displaced the Earth from the center of the Solar System and made it a minor object that orbited the Sun. Over time, our sense of what the Universe is grew larger. By the early 20th century, the Sun was further removed from the center of our collective system of stars that made up the Galaxy. Furthermore distant spiral nebulae appeared to be island "universes" full of stars not unlike our own Milky Way. With the idea of inflation, even our universe has become one tiny fragment of the actual **Universe** that is enormously larger than we can observe. Finally at the dawn of the 21st century, not only do we learn that the baryonic matter that makes us up is only a fraction of all matter, but matter itself is a minor component in the Universe when compared to the even more mysterious dark energy.

If there is any lesson to be learned from this, "Expect surprises" might be an appropriate one. Astronomy has continued to reveal how much more startling and amazing nature is around us than we can possibly imagine. Despite the immense amount of knowledge that we have already gained about the cosmos, it will be a long time before we run out of interesting things to discover.

Appendix A

Notation

A.1 Scientific Notation

Because science often deals with extremely large or incredibly small numbers, a shorthand notation for writing such figures has been developed. First we write an **exponent** as a superscript after a number, which signifies the number of factors of that number to be multiplied together. Thus, an exponent of 2 over a 10 means two tens multiplied together. Here is a list of exponents up to 8:

$$
\begin{aligned}
10 = 10 &= 10^1 \\
100 = 10 \times 10 &= 10^2 \\
1000 = 10 \times 10 \times 10 &= 10^3 \\
10,000 = 10 \times 10 \times 10 \times 10 &= 10^4 \\
100,000 = 10 \times 10 \times 10 \times 10 \times 10 &= 10^5 \\
1,000,000 = 10 \times 10 \times 10 \times 10 \times 10 \times 10 &= 10^6 \\
10,000,000 = 10 \times 10 \times 10 \times 10 \times 10 \times 10 \times 10 &= 10^7 \\
100,000,000 = 10 \times 10 \times 10 \times 10 \times 10 \times 10 \times 10 \times 10 &= 10^8.
\end{aligned}
$$

A simple way to remember what number is represented by the shortened **scientific notation** form is to use the exponent as the total number of zeroes in the "long" version of the number. "One million" of 10^6 will have 6 zeroes, meaning written out, it will be "1,000,000."

We can also define negative exponents, based on the number of factors of 10 that have been divided:

$$\begin{aligned} 0.1 &= 10^{-1} \\ 0.01 &= 10^{-2} \\ 0.001 &= 10^{-3} \\ 0.0001 &= 10^{-4} \\ 0.00001 &= 10^{-5} \\ 0.000001 &= 10^{-6} \\ 0.0000001 &= 10^{-7} \\ 0.00000001 &= 10^{-8}. \end{aligned}$$

Comparing the two sequences, it also makes sense to define:

$$1 = 10^0.$$

We can now use this system to express very large numbers, such as the length of a light year (the distance light travels in one year):

$$1 \text{ lightyear} = 9,460,000,000,000 \text{ kilometers} = 9.46 \times 10^{12} \text{ kilometers}.$$

Very small numbers can also be expressed such as the density of gas out in interstellar space between the stars:

$$\begin{aligned} &0.000000000000000000000002 \text{ grams per cubic centimeter} = \\ &\qquad 2 \times 10^{-24} \text{ grams per cubic centimeter} \end{aligned}$$

A.1.1 "Orders of Magnitude"

If you read enough about astronomy or hang out long enough with scientists, you will inevitably run into or hear the phrase *order of magnitude*. "The mass of Jupiter is *two orders of magnitude* larger than the mass of the Earth." "However the mass of the Sun and a white dwarf are the same order of magnitude." Clearly numbers are being compared here, but what and how exactly?

One interpretation for order of magnitudes is the following: Two values are of the same order of magnitude if they are within a factor of 10 of each other. Thus the numbers 1.0, 2.58374, 7.5, and 9.1 are the same order of magnitude as each other. So are 134, 298.3847, 534.001, and 877. However 3 and 30 are different orders of magnitude: 30 is one order of magnitude larger than 3 since it is 10 times larger. Similarly 4 and 400 are two orders of magnitude difference, since you have to multiply by 10 twice to get from one to the other (or divide by 10 twice going the other way).

A more common (but slightly different) interpretation is that the order of magnitude is to compare the exponents you would need to write the numbers in scientific notation, so that there is only one digit to the left of the decimal place. Thus 102 and 791 can be written as:

$$102 = 1.02 \times 10^2$$
$$791 = 7.91 \times 10^2$$

In both cases, the exponent at the end is 2, so the two values are of the same order of magnitude. Now compare this to the number 15.15:

$$15.15 = 1.515 \times 10^1,$$

which has an exponent of 1. The number 15.15 is 1 order of magnitude smaller than the previous two numbers.

Things get fuzzier when one of the numbers creeps uncomfortably close to the next order of magnitude. Are 10 and 99.99 the same order of magnitude? The second number is still less than 100, which would be a natural cut-off for the first order of magnitude according to the first method. In scientific notation, they are 1.0×10^1 and 9.999×10^1, so they retain the same exponent, which means they are again the same order of magnitude according to the second method. However one can also make the argument that 99.99 rounds up to 100, which would place it 1 order of magnitude away from 10 (via either method). Since the idea of an order or magnitude was originated to loosely compare numbers, there is some arbitrariness to deciding which category a number falls into. So for the case of 10 and 99.99, you might choose to go one way or the other.

One extenuating circumstance may push you to choose one over the next. That is if there is some uncertainty the measurements (and there is *always* some uncertainty in any empirically derived number in science). If the second measurement is not only 99.99, but has error bars of 25, then technically, the number is 99.99 ± 25[1]. The actual value can be 74.99 or as high as 124.99. But because the uncertainty is so large, one cannot realistically give the measurement to that many places of numerical accuracy after the decimal point. It is more honest to say 100 ± 25, keeping the measurement to the same number of significant digits as the error. In this case, 100 ± 25 is two orders of magnitude different from 10.

A.2 Scientific Units

We will use almost all metric—or "SI" for *Système Internationale*—units in this course. For length and distance measurements, we will use the meter (m) and its

[1] "Ninety-nine point nine nine plus-or-minus twenty-five.

variants:

$$\begin{aligned}1 \text{ m} &= 39.37 \text{ inches}\\ &= = 100 \text{ cm} = 10^2 \text{ cm (centimeters)}\\ &= = 1,000 \text{ mm} = 10^3 \text{ mm (millimeters)}\\ &= = 1,000,000,000 \text{ nm} = 10^9 \text{ nm (nanometers).}\end{aligned}$$

For mass, we refer to the gram (g) and kilogram (kg):

$$\begin{aligned}1 \text{ kg} &= 2.20 \text{ pounds}\\ &= 1000 \text{ g} = 10^3 \text{ g.}\end{aligned}$$

The SI unit for temperature is the degree Kelvin (K). It is similar to the Celsius or centigrade degree, so that 1° K = 1° C But while 0° C is set to the freezing point of water, 0° K is defined to be at **absolute zero**, when all thermal motion stops, and therefore the coldest temperature possible. To convert from degrees Kelvin to degrees Celsius:

$$\text{Temp}(^\circ \text{ C}) = \text{Temp}(^\circ\text{K}) - 273.$$

To do the opposite and go from Celsius to the Kelvin scale:

$$\text{Temp}(^\circ \text{ K}) = \text{Temp}(^\circ\text{C}) + 273.$$

According to these formulae, the boiling point of water is then 100° C = 373° K.

A.3 Astronomical Units

Since the space sciences deal with the vast distances in the universe, a number of length measurements have appeared in the astronomical sciences that are used nowhere else in science. Although they have nothing to do with SI units, their use is so widespread that it is unlikely they will go away in our lifetimes or perhaps ever.

The first is the *astronomical unit* or *AU*, and this is defined to be the distance (technically the semi-major axis distance) between the Earth and the Sun. It is defined to be:

$$1 \text{ A.U.} = 149,597.892 \text{ km} = 1.49597892 \times 10^{13} \text{ cm.}$$

You will find it used whenever distances in the Solar System are referred to. You may even see this unit in measurements of other solar systems, or proto-solar systems: a planet orbiting 6 AUs from its parent star, or an accretion disk 60,000 AUs in diameter around a Sun-like star.

A far larger unit of measuring distance is the *light year*, the distance that light, moving at 299,790 km s^{-1}, covers in a single Earth year. This is a unit that is appropriate for describing the distances between stars, and is:

$$1 \text{ ly} = 9.4605 \times 10^{12} \text{ km} = 6.324 \times 10^4 \text{ AU.}$$

Another unit similar to the light year is the *parsec*. It derives from the measurement of distances to stars using the technique of "trigonometric parallax," a topic which we will not have time to delve into in this course. However a parsec (abbreviated as *pc*) is a little over 3 light years:

$$1 \text{ pc} = 3.0856 \times 10^{13} \text{ km} = 2.0626 \times 10^5 \text{ AU} = 3.2616 \text{ ly}.$$

What about distances far larger than a parsec or a light year? As you have seen in Lecture 6, the distances between galaxies and clusters of galaxies can be many millions of light years. Instead of coming up with a new unit for this size scale, astronomers use the SI method of attaching prefixes to existing units. Just as we can scale up a meter by 1000 times and call it a "kilometer," a distance 1000 times a parsec is a *kiloparsec* or *kpc*. A kiloparsec or kilo-light year is appropriate for describing distances from one end of a galaxy to the other. In Fig. 5.22, we see that the visible Milky Way disk is about 50 kpc in radius or 100 kpc in diameter.

For distances between clusters of galaxies, we must resort to the "Mega-" prefix, where a *Megaparsec* (or *Mpc*) is 1 million = 1,000,000 parsecs. Figs. 6.15 and 6.18 could have been re-labeled using Mpc or Mly units. Finally for the scales of voids and filamentary superclusters in the large-scale structure in the observable Universe (such as Fig. 6.21, we can go to the even larger prefix of *Gigaparsecs* (Gpc), which is 1 billion parsecs.

...

One final note about astronomical units concerns units that are based on our Sun. When looking at the energy output of other stars, it is often useful to compare them to that of the Sun. Therefore, you might read or hear that a particular star has 12 $L_\odot$, or 12 times the luminosity of the Sun. (The $\odot$ symbol refers to the Sun.) Similarly one can also use the mass of the Sun, as a unit: you can say a neutron mass has a mass of 2.5 $M_\odot$. Or you can claim a supermassive black hole has a radius of 10 $R_\odot$, or 10 times the Solar radius.

Just for reference, the solar mass, luminosity, and radius can be written in traditional SI units as:

$$\begin{aligned} 1 \text{ M}_\odot &= 1.989 \times 10^{30} \text{ kg}, \\ 1 \text{ L}_\odot &= 3.826 \times 10^{26} \text{ Joules/sec}, \\ 1 \text{ R}_\odot &= 6.9598 \times 10^5 \text{ km}. \end{aligned}$$

Appendix B

The Scientific Process

Astronomy is almost unique amongst the sciences because it is rare for researchers in this field to have direct contact with their subject matter. Except for the instances of meteorites that are found on the Earth, and for the kilograms of rock brought back from the Moon in the Apollo missions, there is nothing to handle, nothing physical to touch or probe, nothing to store in a drawer or to catalog in a museum. Instead the vast majority of the information obtained in this endeavor is via light, whether it is the light that is focused at a telescope on Earth or in orbit around it, or the radiation that is measured and tabulated by the robotic spacecraft sent to the other planets in our Solar System.

Thus it is very much unlike other physical sciences. Chemists and laboratory physicists have the subject under investigation with them in the the lab. Geologists and archaeologists can search for and dig up rocks and artifacts in the field. Afterward they take it home where they can further probe and investigate their samples using the full array of instrumentation available to modern scientists.

Furthermore astronomy is also amongst a minority of sciences where observation is the only way to obtain data. One cannot set up an experiment involving the planets or their moons, arbitrarily picking atmospheres and distances from the Sun. We cannot arrange for a cluster of stars with desired characteristics, and wait around for several hundred million years to see them evolve. In some ways, the astronomical sciences are similar to paleontology and evolutionary biology. The time scales of the space sciences, and of the rates of speciation and change in genomes in biology are much too long to study in "real" time. In all of these disciplines, we do not yet (or perhaps never will) have enough control over the environment to perform controlled experiments.

The **scientific method**, which you may remember learning about in grade school, is usually summarized as:

1. Form a **hypothesis** about a phenomenon that you are interested in investigating.

2. Design an **experiment** that will test this hypothesis. Run your experiment and carefully tabulate your **observations**.

3. Do your results match the predictions that your hypothesis would give? If not, go back and repeat Step 1 above, and repeat these steps.

Except for a very few rare cases, there is no way in the space sciences to create or run experiments on objects that extremely far away, or to study phenomena that change on timescales far longer than our own lives. If experiments are not possible, how does a research astronomer or planetary scientist[1] do their work?

Without objects at hand to study, astronomers are forced to make as many observations as possible, and then to come up with models that can coherently describe these observations. For instance, in the planetary sciences, we make as many detailed observations of the terrestrial planets as possible from telescopes here on Earth as well as from spacecraft probes. These data include the compositions of their atmospheres, the minerals that make up their surfaces, and counting the number of craters for some surfaces. Then using everything that know and understand about geological and atmospheric processes on the Earth, as well as fundamental chemistry and physics from laboratory studies, we are able to make **inferences** or informed guesses as to what is occurring on distant planets.

Based on the volume of observations, scientists can construct a hypothesis. Notice that although this is the same as Step 1 above, it actually occurs *after* the data-taking or observations! With this data, scientists can also construct elaborate computer *simulations*. These simulations or computer models can be used to directly test the observations. Given what we know about how close Venus is to the Sun, how much heating it receives from solar radiation, what its atmospheric composition including the total amount of greenhouse gases, can we construct a computer model which will tell us what the temperature at the surface of the planet will be? Even more importantly, can we make *predictions* that can be tested with new data that we do not currently have? If the theoretical predictions are confirmed by new observations, then we have additional evidence that the model is correct (or at least more correct than competing models that do not have as much observational support).

Similarly in the field of stellar evolution, we have never studied a star, following it from its inception to its death. However we can study thousands of stars of different masses and at different stages in their lives. Using careful observations of the spectra

[1]We use the term *planetary scientist* to refer to anyone whose research involves objects inside the Solar System, and the term *astrophysicist* to signify someone whose primary research is of objects outside the Solar System. The term *astronomer* can be used to signify either, although it tends to be applied to astrophysicists more than planetary scientists. These are distinctions that astrophysicists and planetary scientists make between themselves, although it can be confusing to laypersons at first, because *astronomy* is generically used to refer to all the space sciences, regardless of whether it is in or Solar System or not, and the term *planetary sciences* is not widely used outside of the astronomical community.

from these stars (including the absorption lines arising from the stellar atmospheres), increasingly more detailed computer simulations of entire stars have been made which accurately predict not only the various stages of stellar evolution but also the exact quantity and variety of spectral lines one would expect to detect.

The scientific process for astronomers is a bit messier than the original scheme outlined:

1. Make as many detailed observations as possible of the objects or phenomenon you are investigating.

2. Form a *hypothesis* that will explain the observational data that you have collected.

3. If possible, also test the hypothesis by running computer models or simulations of the phenomenon.

4. At a minimum, your hypothesis should be able to explain all existing observations. Are there any contradictions or inconsistencies? How can these be accounted for? If not, how must the hypothesis be modified to resolve these discrepancies?

5. With the new hypothesis, go back to Step 1, take more data, and repeat this procedure while continuing to re-work your hypothesis.

If it seems to you that the scientific process implies that the process never quite stops, you are correct. New observations and additional data are always welcome to test a current hypothesis. This is especially true in the astronomical sciences where every single observation and scrap of data has to be carefully saved, analyzed, and weighed against the prevailing hypothesis. The observational data is usually very limited, and there is never enough information for us to make a conclusive statement one way or another. However we can say with varying degrees of certainty the truthfulness of different hypotheses based on the available data. As we trust our hypothesis more with the build-up of evidence supporting it, competing hypotheses lose their luster. A hypothesis that has a solid record of support from observations slowly turns into a **theory**.

A theory in the sciences is not just a guess, a conjecture, or other similar synonym from our everyday speech. To scientists, a theory is an idea that has accumulated a vast wealth of supporting detail, and is generally agreed to be true. There is *always* a small chance that a theory can be proved wrong and overturned, but as the supporting evidence for it builds up, this becomes smaller over time. Even if a piece of contradictory datum is found, the theory is not junked. We must ask how reliable is the new information? What are the error bars associated with the measurement? How does it stack up with the data that *support* the theory? Can the theory be modified to explain the discrepant evidence? And if enough new facts come to light so that the theory *has* to be thrown out, why did all of the previous evidence support the original theory?

Can we eventually get to the point where we have absolute certainty about a theory? The answer, I am afraid, is never. We may be 99% sure, or perhaps even 99.9999% sure, but there is *always* some amount of uncertainty in the sciences. By definition, science is involved in investigating nature, and no matter how carefully you make your observations or collect your data, there is always the off chance that what you have tabulated is incorrect. As a result, *because* science is ultimately based on observed reality (despite the fact that theoretical work, computer simulations and modeling are all important), then if your observations are suspect, then the conclusions from them are suspect as well.

In mathematics, one can start with a set of **axioms** and derive via a series of logical steps in a **proof** a new set of conclusions. For instance, using the set of Euclidean axioms in geometry, one can show that the angles inside a triangle will always add up to 180°, or that given a point and a line, there is exactly one line that can be drawn through the point that will be parallel to the first line.

However the sciences are not mathematics (nor philosophy, for that matter, another branch of knowledge where proofs are used). Although we can use the same deductive procedures that mathematicians use, we cannot be as certain of our conclusions as they are. Our starting statements and observations do not have the same level of certainty that mathematical axioms do. As a result, our theories and hypotheses are constantly undergoing revision as new data becomes available.

As an example, Kepler's three Laws of Planetary Motions were found to accurately explain the motions of the planets in the sky (see §2.2). However science did not stop with Kepler. Kepler's Laws did not adequately explain the motion of the Moon around the Earth, nor the Jovian satellites (once Galileo discovered them) around Jupiter. However an alternative to Kepler's Laws appeared in the 17th century: Isaac Newton (1642–1727) asked *why* did Kepler's Laws work so well, and was able to derive his own theory of motion that not only explained the motions of all the objects in the heavens, but also of objects here on Earth. His ideas of what is now known as **Newtonian mechanics** were not only an improvement over Kepler's Laws, but they also explained *why* Kepler's Laws worked so well for the planets around the Sun. Kepler's work turned out to be just a subset in the grander scheme of Newton's theory.

Physics did not stop there. With increasingly accurate measurements of planetary motions, there were discrepancies in the observed versus predicted positions of Mercury around the Sun. This and other problems opened the door to Albert Einstein's idea of General Relativity (a follow-up to the earlier theory of Special Relativity). It not only had the explanatory power of Newtonian mechanics, but covered new phenomena (such as the odd behavior of Mercury), and made additional predictions that could be used to test the theory. As a result of its success in passing this and other observational tests, Einstein's Theories of Special and General Relativity are the best descriptions we have of explaining motion in the universe.

You might also notice that successful old theories do not entirely disappear (unless they are so horribly wrong that they have no application in modern life). We can still use Kepler's Laws today to describe the motions of the planets, as long as we confine it *only* to the objects orbiting the Sun, and we do not need the results to a high level of accuracy. Newtonian mechanics is still perfectly acceptable for calculating the positions of most orbiting bodies in the Solar System, and is all that is necessary for accurately plotting the trajectories of spacecraft like Cassini to Saturn.

Appendix C

Electromagnetic Radiation

Because the stars are so far away—even the nearest star, the Sun is 149,600 kilometers from the Earth—we can only learn about them by observing and recording the light that they emit. To understand how we learn about the physical attributes of stars, and intuit the processes by which they are born and die, we must understand the basic processes by which light is emitted by atoms. This Appendix is therefore devoted to fundamental atomic physics and electromagnetic radiation.

C.1 Electromagnetic Radiation

The term for the majority of the emission from astronomical objects that we detect here on the Earth is **electromagnetic radiation**. The radiation is energy that is transmitted across space, without the need for an intervening medium. The reason why it is "electromagnetic" is that it consists of fluctuating electric and magnetic fields. **Visible light**, which our eyes are sensitive to, is only one part of the **electromagnetic spectrum**. **Infrared**, **ultraviolet**, **microwave**, **radio**, **ultraviolet**, **x-ray**, and **gamma rays** describe other portions of the electromagnetic spectrum, although they are invisible to our eyes. These terms all describe electromagnetic radiation, but these different phrases exist because historically, it took physicists many decades to realize the different phenomena they were studying were merely different aspects of the same thing. As a result, we will use the word "light" generically to mean "electromagnetic radiation."

Another synonym for electromagnetic radiation is electromagnetic **waves**. This is because light is a wave phenomena, similar to waves moving across a pond or sound waves traveling through the air. A typical wave consists of several different aspects: the **crests** at the tops of the wave, the **troughs** at the bottom portion of the wave, the **wavelength** which measures the distance from trough-to-trough (or equivalently from crest-to-crest). The **amplitude** is the height of the crest above, or the depth of the trough below the undisturbed neutral state of the wave. A wave will also have a direction of motion, where energy is transfered from one place to another along this direction. The wave will also have a **wave speed**. If we stand at a stationary point and watch a wave with a fixed speed pass by, we can count the number of crests (or troughs) flow by within a fixed period of time. The measure of the number of wave **cycles** per second is the wave's **frequency**.

One way to create a set of waves is to start with a perfectly still pond, and then to toss a pebble into its center. The energy of the falling pebble impacting the surface of the pond is transformed into the energy in the waves, which propagate outward, in concentric circles. The waves in the water have all the characteristics described above. The amplitude we can measure if we place a cork in the water, and we watch how much it bobs up and down as the waves pass. The wavelength is simply the distance from the crest of one wave to the next.

Similarly if we can create electromagnetic waves by disturbing an electric charge, such as the oscillating electrical currents in a radio or TV antenna. Electromagnetic waves propagate outward carrying along energy with them. There are two important differences between electromagnetic waves and water waves. First, although the waves at the surface of a pond are actually propagating through the medium of water, electromagnetic waves do not need a medium to travel through. Although light can move through the medium of air, it can move through a perfect vacuum as well. Second the different forms of electromagnetic radiation all travel at exactly the same speed. This is the **speed of light**, which in a vacuum is 299,792 kilometers per second[1].

The speed of light is the absolute speed limit in our universe. Nothing can go faster than light. Any signal that we receive from the elsewhere in the universe or send out from the Earth will travel at best at the speed of light. Any other method of transmitting information will be slower.

Although light travels extremely fast by our standards, it moves at a *finite* rate, although on human scales, it does seem almost infinitely fast. Because objects in astronomy are so distant, there is always a delay between when an event occurs and when we can observe it. Even in our Solar System, we have a 3 second delay between the Earth and our Moon, an 8 minute delay for a signal to get from the Sun to the Earth, and up to 20 minutes between the Spirit and Opportunity rovers on the surface of Mars and their controllers on Earth. Outside the Solar System, the distances become far vaster; light takes 4 years to travel from the nearest star, Alpha Centauri to the Earth. To get to the Andromeda Galaxy, the nearest large galaxy of stars outside of our own Milky Way, takes 1.5 million years.

Because electromagnetic waves all travel at the same speed, the only way they can be different is in their wavelengths. The wavelength (or equivalently the frequency) determines the intrinsic "color" of the light. For the visible spectrum, we are familiar with the colors red, orange, yellow, green, blue, and violet. Each of these colors have a different wavelength, from about 0.0000007 meters (or $7.0 \times 0.0000001 = 7.0 \times 10^{-7}$ meters in scientific notation) for red light to 0.0000004 meters (or $4.0 \times$

[1]In a physical medium such as air or water, the speed of light is slower.

10^{-7} meters) for violet light. Since it is cumbersome to write the wavelengths of visible light using meters as a unit, another useful unit is the **nanometer** (nm), which is one billionth or 10^{-9} of a meter. The wavelength of red light is 700 nm, and blue light is 400 nm.

The colors of the visible part of the spectrum can be present in light even when we cannot see them. For instance sunlight looks white to our eyes, but when we send the sunlight through a prism, the light will be broken up into the constituent visible colors. We can use an even more sophisticated instrument to study the spectrum or the different wavelengths of light. Such a **spectrograph** can be built to spread out light from virtually every domain of the electromagnetic spectrum, from the radio to X-rays.[2]

When light has a wavelength that is longer than red light or shorter than violet light, then this electromagnetic radiation falls outside the range of visible light. Depending on the exact wavelength, the light is either infrared or radio if the wavelength is longer (or a smaller frequency). Radio waves include microwaves, and the radio frequencies used by AM and FM radio, TV, cell phones, and radar. We cannot see infrared light, but we can sense it with our skin as heat. The light that has wavelengths just shorter than violet is ultraviolet, and is the type of radiation responsible for sunburns and suntans. At even shorter wavelengths (greater frequencies) there are X-rays, known to us for their ability to pass through human tissue and hence its use in medicine. The shortest wavelength radiation are the gamma rays, which are associated with radiation emitted during some nuclear reactions and certain types of radioactive decay.

C.2 Spectral Lines

Atoms can be made to both **emit** and **absorb** light at very specific wavelengths. We can approximate the structure of an atom with a very simple model of a **nucleus**, consisting of positively charged **protons** and neutral **neutrons**. Around the nucleus is a cloud of negatively charged **electrons**. In our simplified view, the electrons occupy levels around the nucleus, the **electron orbitals**[3], analogous to the planets in orbit around the Sun. Ordinarily the electrons are found in **ground states** at fixed distances from the nucleus. Unlike planets around the Sun, these orbitals have **quantized** energies, meaning their distances from the nucleus are fixed. Furthermore each orbital has a maximum number of spaces for electrons. Once an inner orbital is filled, additional electrons can only be located at orbitals further from the nucleus.

---

[2]Gamma rays and X-rays are actually better studied with "photon" counting devices.

[3]The exact locations of the electrons around the nucleus and the way they move are described by the theory of *quantum mechanics*. Our simplified theory of the atom is actually far from correct, when compared with the quantum mechanical view. However the "correct" model of the atom is beyond the scope of this course, so we will continue to use the simplified view.

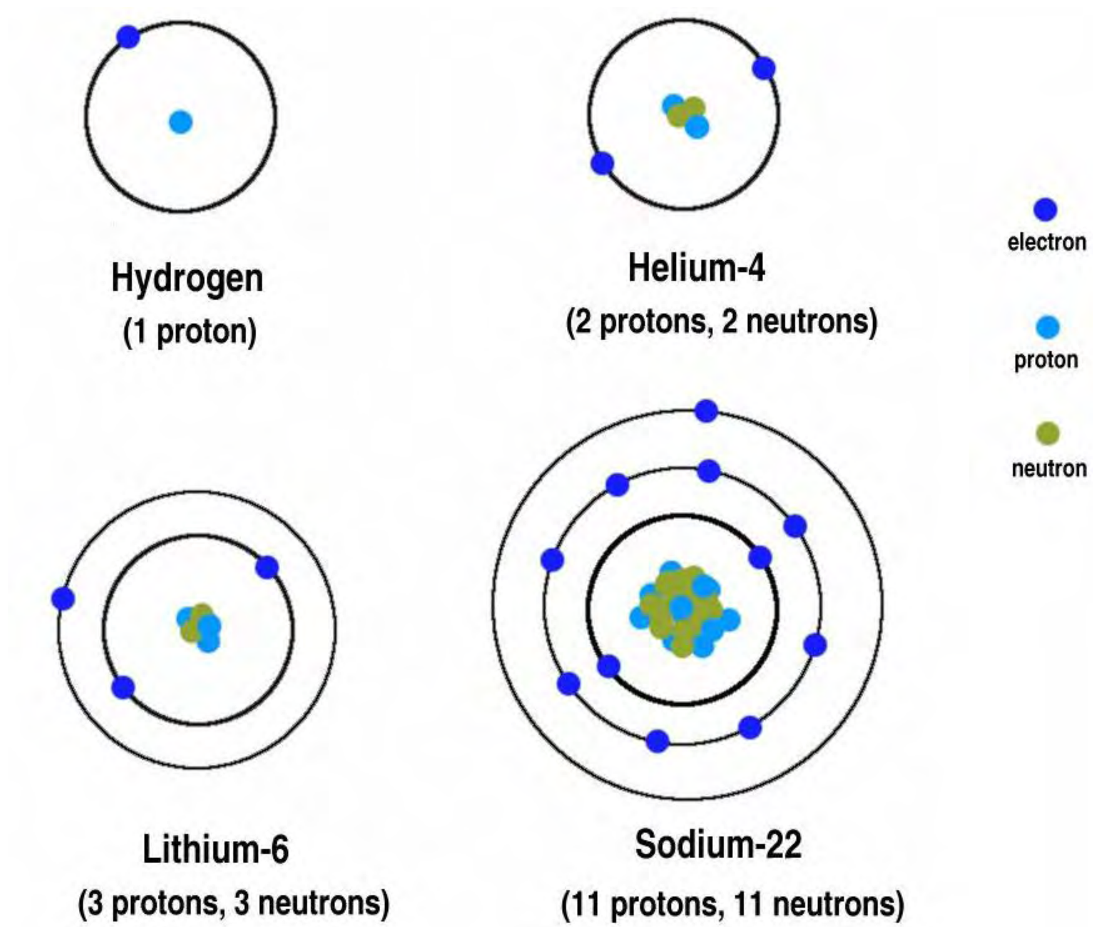


Figure C.1: The Bohr Model of the Atom

Our simplified model of the atom shows protons and neutrons located in the nucleus, while elecxtrons stay in orbitals of a fixed radius from the center. Only a finite number of electrons can fill each orbital; for instance, the innermost orbital holds only two electrons, and once that is filled, the next electron goes into the next orbital. The number of electrons matches the number of protons in normal atoms. If an atom is ionized*, it has either lost or gained an extra electron.*

Atoms are normally in a ground state. However an atom can be **excited** if energy is pumped into it which elevates an electron from its original orbital into a higher one. The amount of energy needed to go from one orbital to another is fixed and very specific. There are two ways in which an atom can be excited. It can absorb electromagnetic radiation with the energy equivalent to the difference in energies of the two orbitals. Or the atom can collide with another atom or particle, which deposits the necessary energy. Once excited, the atom cannot remain in its excited state forever. The electron will drop back down into its ground state in about 10^{-8} seconds, and will give up its extra energy, which is emitted as electromagnetic radiation.

Because the energy that can be absorbed and then re-emitted by an atom is highly specific, the electromagnetic energy is also specific as well. The radiation is thus absorbed and emitted in energy "packets" called **photons**. Every photon of light is associated with a particular wavelength or frequency, with a direct correlation between the frequency and the photon's energy. The greater the frequency (shorter the wavelength), the higher the photon energy. An ultraviolet photon therefore packs more energy than an infrared photon. The energy necessary to pump an electron from one orbital to another must be provided by a photon with the exact energy difference between the two orbitals. Once the electron de-excites back to the ground state, a photon with that same energy is released.

We have been speaking of photons as if they were actually like particles, rather than waves. It is another consequence of quantum theory that light turns out to have both particle- and wave-like attributes. We will still describe light using wave-like descriptions, such as the terms "frequency" and "wavelength." But in other situations, it is sometimes easier to describe light as arriving as discrete photon packets, almost like small BBs. Light will have both wave and particle-like characteristics depending on the context, and we will use both depictions here.

Each atom of a different element will have a very different set of electron orbitals with energy differences found only for that element. Thus an atom of an element can **absorb** light at wavelengths that are specific to that element, and once excited, it can **emit** light at those same wavelengths as well. If we use a prism to spread out the colors of the light emitted from an atom, the specific wavelengths of emitted light will show up as **spectral lines**. If we carefully map the wavelengths of light emitted from excited atoms of every element we can use this information to determine the composition of atoms that are too far away for us to investigate.

If sufficient energy is applied to an electron around an atom, it may be knocked completely out of the atom's grasp. An atom losing one or more of its electrons is in an **ionized** state. Because opposite charges attract, the electrons close to the nucleus in the inner orbitals are held much more tightly than electrons in far orbitals. A light atom like hydrogen has one proton in its nucleus, so the attractive strength for its electrons is not very large. However an electron in the innermost orbital is very close to the nucleus, so the electron is bound tight enough that it requires an

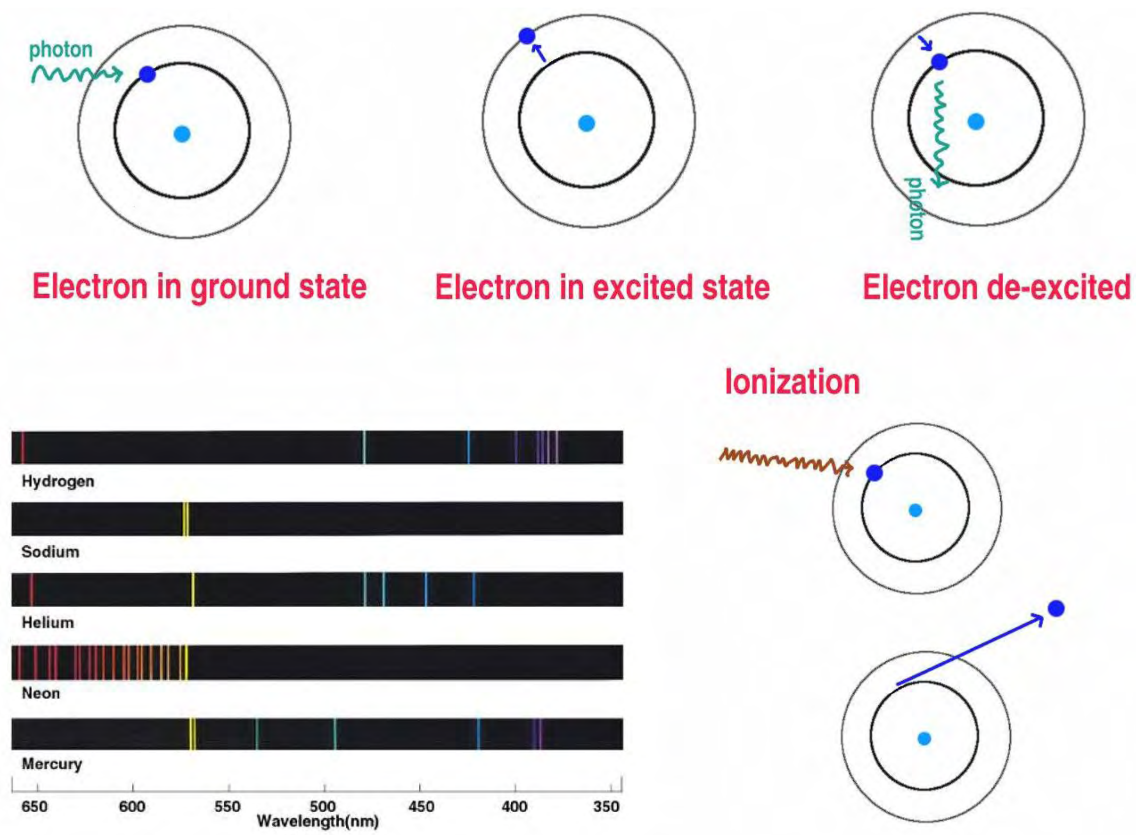


Figure C.2: Electron Excitation

An electron at rest in its ground state *can be excited by an input of energy, in this case an electromagnetic photon* (left)*; a photon with the energy equivalent to the energy difference between the two orbitals causes the electron to jump to a higher orbital. When the electron* de-excites*, it drops back down to the original orbital, and releases its energy in the form of another photon* (right)*. If a photon has too much energy, it can cause the electron to escape entirely from the atom* (bottom right)*. The atom has now been* ionized*. It is this emission of photons with highly specified energies (or equivalently wavelengths) that causes every atom to have a unique spectral signature* (bottom left)*.*

ultraviolet photon to ionize a hydrogen atom. For a heavier atom like uranium, there are far more protons in the nucleus (92 in this case), so the inner electrons are far more tightly bound than in hydrogen. However the outermost electrons are further away, and hence more loosely bound. Heavy elements are therefore easy to ionize by photons in the visible part of the spectrum, at least for the first several electrons (how many will depend on the atom). Extracting the innermost electrons from these atoms will require energetic X-ray photons.

C.3 Blackbody Radiation

Our discussion of the emission and absorption of light only refers to free solitary atoms, such as would be found in a gas. However most atoms we are familiar with are bound up in molecules, and these molecules are found close together in the form of solids or liquids. Atoms in such a state will still emit and absorb energy at wavelengths specific to their electron orbital energy levels. But the close proximity of other atoms means that any emitted radiation is not free to escape but is immediately re-absorbed. The negatively charged electron clouds of neighboring atoms can also alter and change the orbitals slightly. The net effect of this for a material solid is that radiation emitted from a solid substance as a whole has a **continuous spectrum** instead of individual spectral lines. Colors of all wavelengths are being emitted, although the **intensity** of each color is different. The total amount of emission depends on the **temperature** of the material. The higher the temperature, the more energy the solid will contain, and the more emission will show up.

Substances at room temperature do not have enough energy to emit any visible light at all; most of their radiation is in the form of infrared energy. Once you heat a object to hundreds or thousands of degrees Celsius, there is enough energy present for the object to glow at visible wavelengths. This is the reason why the electric burner on our stoves glow faint red when we first turn them on, and then will have a bright orange-ish red emission once they are hot. Objects at increasingly hotter temperatures will start emitting more light at shorter wavelengths—more greens and blues—which when added to the red and yellows result in a white hot light. Objects at 10,000° Celsius or more will start emitting more ultraviolet radiation. If you raise the temperature of a solid to millions of degrees, X-ray emission will start to become important.

The amount of radiation emitted at each wavelength from a solid is described by the **blackbody** curve. Electromagnetic radiation from solids is hence called **blackbody radiation**. The curve peaks at a specific wavelength, and falls off on either side. The curve is not symmetric, so that the falloff at higher frequency (shorter wavelengths and hence greater energy photons) is steeper than the falloff at the longer wavelength end. The shape of the curve is the same regardless of the temperature. However increasing the temperature of an object will increase its overall intensity of emitted light, and also shift the peak of the curve to the shorter wavelength side.

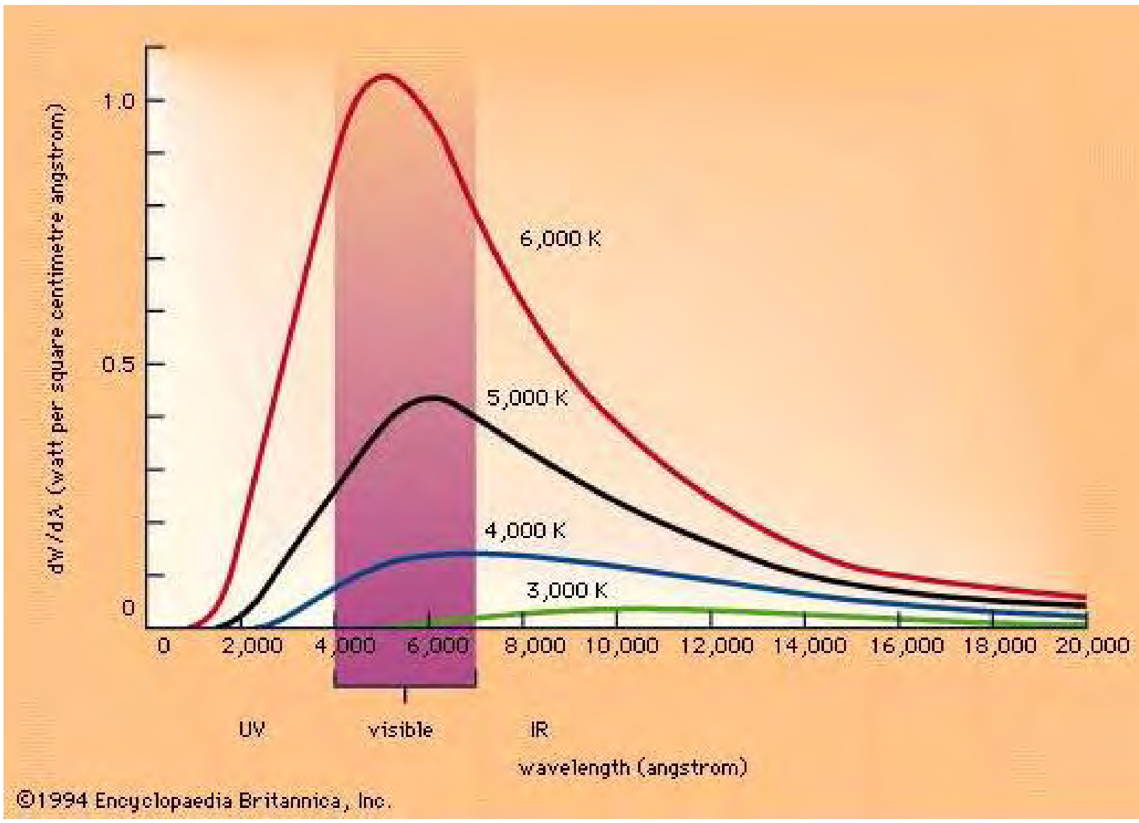


Figure C.3: Blackbody Spectra at Different Temperatures

Blackbody spectra from objects with a range of temperatures, from 3000° K to 6000° K. As the temperature of a solid object goes up, the total emission increases, while the peak of the emission shifts to bluer (higher energy) wavelengths. For a body radiating at 3000° K, most of the emission is still in the infrared portion of the spectrum, although there is some red visible light as well. For the 6000° K object, the most emission comes in blue-ish green, although there is plenty of light from other visible wavelengths as well.

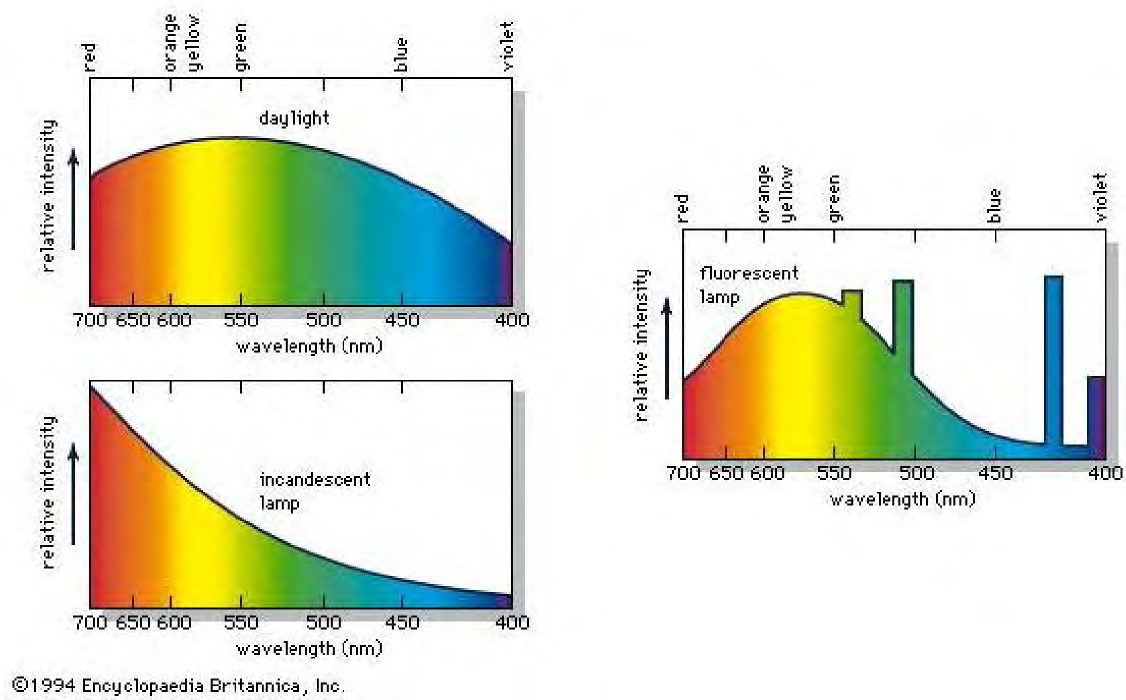


Figure C.4: A Variety of Blackbody Emitters

Three examples of blackbody emission in the narrow visible light part of the electromagnetic spectrum. (Top left) *Daylight comes from the Sun, a dense, hot mass of gas that emits like a solid object. Its emission peaks in the yellow-ish green.* (Bottom left) *An incandescent lamp glows from a solid tungsten filament that has electricity running through it. The temperature of the lamp is much lower than that of te Sun, so most of its visible emission is in the red (although its actual peak of emission is in the infrared).* (Right) *A fluorescent lamp is a glass tube filled with a gas whose atoms are excited by electricity coursing through the lamp. The interior surface of the tube is also filled with a fine solid powder. A spectrum of a fluorescent lamp therefore contains emission spectral lines from the gas, as well as a blackbody curve from the powder.*

This shift in the curve thus corresponds to the same shift that we see in the electric burner as it moves from red hot to orange hot.

Further Reading

The following is a list of books to learn more about topics of this course. The order however is not alphabetical, but in the rough order of the lecture subject matter. For instance, references for stargazing and star names appear first, while books on the Big Bang are last. A brief description appears where I have comments. The first three titles are books that try to cover much of the history of the Universe in one shot.

Terence Dickinson, *The Universe and Beyond*, 4th edition, 2004, Buffalo, NY: Firefly Books. [This book reviews most of the astronomy from the Solar System (its weakest coverage) onward to the edge of the observable universe and beyond. It hits the themes of enormous scales that we have touched upon in class, as well as issues of astrobiology and life elsewhere, which we never got around to. The pictures are well chosen, and the text is current to boot.]

Out of print Carl Sagan, *Cosmos*, 1980, New York: Random House. [This companion to the PBS television series is a somewhat dated work, but it does a superb job of showing the history, the science, as well as the surprising romance and passion one can have in studying the Universe. Look for a used copy of the trade paperback or hardcover, since the illustrations are integral to the book, and they are missing in the paperback that is currently in print.]

Neil DeGrasse Tyson & Donald Goldsmith, *Origins: Fourteen Billion Years of Cosmic Evolution*, 2004, W. W. Norton & Company.

Richard Hinckley Allen, *Star Names: Their Lore and Meaning*, 1963, New York, NY: Dover Publications. [A reprint of a title originally published in 1899, this is one of the best books for etymologies and multi-cultural lore on stars and constellations. References go back to the Greeks, Arabs, Chinese, Romans, Egyptians, Norse, and more. As a result, it is more of a reference work for the modern reader than for literary enjoyment.]

J. L. E. Dreyer, *A History of Astronomy from Thales to Kepler*, 1953, New York, NY: Dover Publications. [More than you would ever want to know concerning early ideas of astronomy from the ancient Greeks on till the 16th century.]

Rajiv Gupta (Editor), *Observer's Handbook 2004*, 2004, University of Toronto Press. [A new edition appears each year, so look for the 2005 edition soon. This contains listings of all important astronomical phenomena that will appear in the sky for the upcoming year. Useful for any serious amateur astronomer.]

Jay M. Pasachoff & Roger Tory Peterson, *A Field Guide to the Stars and Planets, 4th Edition*, 1999, Boston, MA: Houghton Mifflin. [A field guide with star charts, notes on important objects, history and lore concerning the constellations, an atlas of the Moon, and charts and tables filled with countless other information in a very compact form.]

Ian Ridpath & Wil Tirion, *Stars & Planets, 3rd Edition*, 2001, Princeton, NJ: Princeton University Press [Part of the Princeton Field Guide series, this is a another good all-around guide to the sky.]

Guy Consolmagno, Dan M. Davis, Karen Kotash Sepp, Anne Drogin, & Mary Lynn Skirvin, *Turn Left at Orion*, 2003, Cambridge University Press [One of the most useful guides to finding objects with your telescope for amateurs, especially if your telescope lacks the ability to automatically go to an object you've selected. With a basic knowledge of the major constellations, you can learn to ''star hop'' to find obscure objects. Diagrams show what you would actually see in an eyepiece.]

Chet Raymo, *An Intimate Look at the Night Sky*, 2003, Walker & Company [Part star atlas, and part meditation on astronomy and the night sky. Each of the twelve chapters corresponds to a month of the year and opens up with a star chart pointing out objects and constellations of interest. The essays however cover all of the important astronomy topics with wonderfully lyrical writing.]

J. Kelly Beatty, Carolyn Collins Petersen, & Andrew Chaikin, *The New Solar System, 4th Edition*, 1998, Cambridge, UK: Cambridge University Press. [Almost the standard reference at the intermediate-level for the Solar System. If you want to fortify your learning with more math and physics, this is the place to come to. It covers most everything of interest, and is illustrated with the latest (as of 1999) spacecraft images. Some portions do get technical, but the book is well worth the money.]

Michael Benson, *Beyond: Visions of the Interplanetary Probes*, 2003, Harry N. Abrams. [Beautiful coffee table book filled with glorious, often-full colored images taken from space of planetary and moon surfaces. There is not much in the way of text, except for somewhat minimal figure captions, and a couple of essays. However it is clear that the incredible pictures are the main point of the book.]

William K. Hartmann, *The Grand Tour: A Traveller's Guide to the Solar System*, 1993, Workman Publishing. [The information in this work is not so much wrong as it is dated. However the paintings by Hartmann, which cover all the planets, major moons, as well as some minor bodies, are very nice.]

Paul Hodge, *Higher Than Everest*, 2001, Cambridge University Press. [Yet another tour of the Solar System, this time, concentrating on locales that would be ideal for the extreme sports enthusiast of the future.]

Kenneth R. Lang, *The Cambridge Guide to the Solar System*, 2003, Cambridge University Press. [A big, coffee table book, which covers some history as well copious information on all of the important objects, atmospheres, surface features, and other miscellania. This is a good comprehensive reference for the novice.]

Understanding the New Solar System, by the editors of *Scientific American*, 2002, Warner Books.

David H. Grinspoon, *Venus Revealed: A New Look Belw the Clouds of Our Mysterious Twin Planet*, 1998, Perseus Books Group. [Local scientist and author provides a comprehensive look at Venus, covering the planet in myth and culture, as well as the latest science. There is also some interesting speculation on the possibilities of past or present life on this greenhouse world.]

Richard Fortey, *Earth: An Intimate History*, 2004, Knopf.

Edmond A. Mathez & James D. *The Earth Machine: The Science of a Dynamic Planet*, 2004, Columbia University Press.

William K. Hartmann, *A Traveler's Guide to Mars: The Mysterious Landscapes of the Red Planet*, 2003, Workman Publishing. [Written literally like a travel guide, the book is broken into chapters that cover distinct regions of the planet. There are plenty of recent pictures taken by orbiting spacecraft.]

Oliver Morton, *Mapping Mars: Science, Imagination, and the Birth of a World*, 2002, Picador.

David A. Rothery, *Satellites of the Outer Planets: Worlds in Their Own Right*, 1999, Oxford University Press. [Covers 18 of the small icy worlds in the outer Solar System, which are large enough to have interesting geology and surface features. Make sure you get the recent (5-year old) edition, which has been updated with images from the Galileo spacecraft.]

Ralph Lorenz & Jacqueline Mitton, *Lifting Titan's Veil: Exploring the Giant Moon of Saturn*, 2002, Cambridge University Press.

John Davies, *Beyond Pluto*, 1997, Cambridge, UK: Cambridge University Press.

Leon Golub & Jay M. Pasachoff, *Nearest Star: The Surprising Science of Our Sun*, 2002, Harvard University Press.

Jack B. Zirker, *Journey from the Center of the Sun*, 2001, Princeton University Press. [A medium level of difficulty book containing more than you might possibly want to know about the Sun, covering the interior as well as the components of atmosphere, convection cells, magnetic fields, flares, hot spots and sunspots. The light paperback discusses many recent observations and current theories.]

C. Robert O'Dell, *The Orion Nebula: Where Stars Are Born*, 2003, Belknap Press. [One of the few books that I know of that has been written for the layperson with an almost exclusive focus on star formation. The subject is the Orion Nebula, the closest H II region and site of OB stars to the Sun, so it does try to cover all aspects of the star formation process.]

James B. Kaler, *Stars and Their Spectra: An Introduction to the Spectral Sequence*, 1997, Cambridge University Press. [Learn about how spectroscopy is used to determine the fundamental properties of stars, as well as covering all the major spectral classes, and evolutionary categories of stars.]

James B. Kaler, *Extreme Stars*, 2001, Cambridge University Press.

James B. Kaler, *The Hundred Greatest Stars*, 2002, Springer-Verlag. [A ''Top 100'' list that tries to capture the full diversity of stellar phenomena in the universe. Each star is shown either by a close-up image or its location in its constellation, and has information for why it is significant.]

Sun Kwok, *Cosmic Butterflies*, 2001, Cambridge University Press. [The ''butterflys'' in the title actually refer to planetary nebulae. The book includes the images returned from the Hubble Space Telescope of some of the most spectacular looking examples. The book details the post-main sequence evolution of low mass stars, as well as the importance of planetary nebulae in the life cycle of stars.]

Laurence Marschall, *The Supernova Story*, 1994, Princeton University Press.

Mitch Begelman, *Turn Right at Orion: Travels Through the Cosmos*, 2001, Perseus Publishing. [Learn astronomy and physics in novel form! Follow the adventures of the narrator who visits stars at different stages in their lives, including red giants, neutron stars, and various black holes, before heading off to the Virgo Cluster with a flyby of the Magellanic Clouds.]

Out of print Nigel Henbest & Heather Couper, *The Guide to the Galaxy*, 1994, Cambridge, UK: Cambridge University Press. [A compendium of the contents of our Milky Way, starting with the local solar neighborhood, and moving outward to ever larger scales, until we cover molecular clouds, supernova remnants and superbubbles, as well as the Galactic Center. The artists' depictions of the different components of our Galaxy are extremely well done.]

Malcolm S. Longair, *Our Evolving Universe, 2nd Edition*, 1996, Cambridge University Press.

William H. Waller & Paul W. Hodge, *Galaxies and the Cosmic Frontier*, 2003, Harvard University Press. [Covers all aspects of galaxies, from morphologies to composition to formation and evolution. The subject of early chapters concentrate on the Milky Way, and then gradually move focus to the Magellanic Clouds, the Local Group, and so on until we get to the superclusters and large scale structures in the Universe.]

Brian Greene, *The Elegant Universe: Superstrings, Hidden Dimensions, ad the Quest for the Ultimate Theory*, 2000, Vintage.

Brian Greene, *The Fabric of the Cosmos: Space, Time, and the Texture of Reality*, 2004, Knopf.

Alan H. Guth, *The Inflationary Universe: The Quest for a New Theory of Cosmic Origins*, 1998, Perseus Books Group.

Robert P. Kirshner *The Extravagant Universe: Exploding Stars, Dark Energy, and the Accelerating Cosmos*, 2002, Princeton University Press. [A whole book devoted to the evidence for dark energy, starting from Einstein's first work on the topic, up to the latest research today.]

Kip Thorne, *Black Holes & Time Warps: Einstein's Outrageous Legacy*, 1995, W. W. Norton & Company. [An excellent mid-level introduction to black holes, evidence for them in the Universe, and their many strange properties. Excellent coverage of a topic that was mentioned just once in class!]

Greg Laughlin, *The Five Ages of the Universe: Inside the Physics of Eternity*, 1999, Free Press. [A history of the Universe starting from the Big Bang and ending at a staggering 10^{100} years later. Lots of interesting ‘‘what if’’ situations.]