The Digital Universe Introduction to Astronomy

COLLABORATORS			
	TITLE : The Digital Universe Introduction to Astronomy		
ACTION	*NAME*	*DATE*	*SIGNATURE*
WRITTEN BY		October 22, 2024	

REVISION HISTORY			
NUMBER	DATE	DESCRIPTION	NAME

Contents

Chapter 1

The Digital Universe Introduction to Astronomy

1.1 The Digital Universe Introduction to Astronomy

This hypertext document is meant to provide a brief introduction to astronomy. The intention is not to teach you everything there is to know about the field, but rather to provide an explanation of several concepts which are critical to being able to use "The Digital Universe" to the utmost of its capabilities.

The following Chapters are available:

1.2 A Short History of Astronomy (ancient - A.D. 999)

For countless millennia, our ancestors have looked at the night sky with awe and wonder. They learned to realize that the Sun gave them heat and light, and that by carefully observing the position of the Sun and phase of the Moon they could measure the passage of time. They realized that while the stars seemed to move across the sky throughout the night, they seemed to retain their relative positions and orientations without change. However, it is only comparatively recently that we have begun to develop what might be considered a 'modern' view of astronomy.

Many of the early western astronomers were of Greek ancestry. The philosopher Aristotle felt there were vast differences between the heavens and the Earth. He believed that on Earth, everything was made up of four basic "elements" - earth, air, water, and fire. However, the sky and everything contained within it were composed of a fifth element, which he called aether.

The astronomer Eudoxus believed that the heavens revolved around the Earth, and as a result the Earth was at the center of the universe. An opposing theory was held by Aristarchus, who felt that certain movements of the sky could be best explained if it were assumed that the Earth revolved about the Sun. Unfortunately, few people accepted Aristarchus' theory, claiming that if the Earth were in fact revolving about the Sun they should be able to feel some sensation of movement.

A few hundred years later, the Greek astronomer Hipparchus became very interested in examining the motions of the planets through the heavens, in an attempt to resolve the debate. While doing so, he became familiar with the formations of stars through which the planets moved. One night, he was astonished to find a star in the constellation of Scorpio that had not been there before. Without realizing it, he had discovered a nova.

This discovery caused quite a stir in the astronomical community. It had previously been thought that although the Sun, Moon, and planets moved through the sky, the stars remained fixed in position and forever unchanging. Suddenly, people had to start modifying their theories of the universe.

It occurred to Hipparchus that perhaps these 'new' stars were not all that rare. It was possible that they occurred on a regular basis, but went unnoticed amidst the great masses of stars visible on a clear night. So he then attempted to record the positions and magnitudes of some 850 stars. By 134 B.C., this first star catalogue was prepared.

After creating his catalogue, he studied earlier observations of the stars and noticed another startling phenomenon. It appeared as though all the stars he was able to see had shifted slightly eastward from their earlier observations. It turned out that Hipparchus had discovered an effect now known as precession. Since such a change was not visible over an individual's lifetime, it clearly was a very slow process. Nevertheless, the effect was real and it served to provide yet more proof that the stars were not as 'fixed' in place as they had once seemed.

At around A.D. 150, Ptolemy made many astronomical observations near Alexandria, Egypt. These observations resulted in a 13 volume work entitled "Mathematike Syntaxis" (Mathematical Composition). The work was based a great deal on Hipparchus' studies, and served to strengthen the belief that the Earth was the center of the Universe. It became highly respected among Greek and Arab astronomers, eventually finding its way to the rest of Europe as well. For 1400 years, nobody questioned the Ptolemaic theory.

See also A Short History of Astronomy (1000 - 1799)

1.3 A Short History of Astronomy (1000 - 1799)

Very little had happened since A.D. 150, when Ptolemy had made his astronomical observations and promoted the belief that the Earth was at the center of the Universe. It wasn't for 1400 years before someone had the courage to question the strongly entrenched Ptolemaic theory.

In 1543, Copernicus tried to revive the idea that the Earth and other planets (though not the Moon) moved around the Sun in circular orbits. In his book, "Concerning the Revolutions of the Celestial Spheres", he demonstrated that this concept could be used to explain the motions of the heavenly bodies. Unfortunately, his theory was difficult to prove. It was expected that if the Earth truly moved, stars should appear to change position slightly as the Earth revolved around the Sun. This phenomena, known as parallax, was much too small for astronomers of the time to measure.

Throughout the late 1500s, Tycho Brahe was attempting to detect this expected change in position. He actually believed in neither the heliocentric nor geocentric theories of the universe. In his "Tychonic system", he felt that Mercury, Venus, Mars, Jupiter, and Saturn (the only planets known in those days) revolved about the Sun, while the Sun (with the planets circling it) revolved about the Earth. In his efforts to determine the true nature of the solar system, he redefined astronomical methods by keeping very accurate and regular records on the positions of the planets and stars, but he too failed to observe the shift of parallax. The controversy between the heliocentric and geocentric theories of the universe remained unresolved.

Johannes Kepler, a German mathematician working with Brahe in the later years of his life, carried on his work after Brahe's death in 1601. It was Brahe's hope that the careful observations he had made of the stars and planets would help Kepler to prove the "Tychonic" theory of the universe. Instead, Kepler's investigations with the data supplied by Brahe suggested that in fact, the heliocentric theory was more appropriate - with one notable modification. All objects were pictured as moving around the Sun (except the Moon, which moved around the Earth), but did so in ellipses instead of the perfect circles that had been previously assumed.

Kepler did more than simply discover that the planets moved in elliptical orbits around the Sun (though that in itself would have been sufficient enough to ensure his place in the annals of history). He developed 3 laws of planetary motion, which have remained indisputed to the present day (other than a few refinements):

1. Every planet follows an elliptical orbit around the Sun. The Sun is located at one of the focal points of the ellipse.

2. An imaginary line drawn between the centre of the Sun and the centre of a planet traces out the same area in a given time, regardless of how far away or how heavy the planet is. A consequence of this is that a planet (and any orbiting body) moves faster when it is nearer to the Sun.

3. The time taken for an object to complete one revolution around the Sun is referred to as its period. The squares of the periods of any two planets is proportional to the cubes of their average distances to the Sun.

In 1608, a Dutch optician named Hans Lippershey invented the first telescope. Unfortunately, he was refused a patent for his invention. The following year, the Italian astronomer Galileo Galilei heard of Lippershey's work and promptly constructed his own. Though his instrument was extremely crude by today's standards, it gave astronomers their first chance to look at the heavens with a device more sensitive than their own eyes.

Galileo used his telescope to advance astronomical knowledge by leaps and bounds. He discovered that the Moon was not a smooth sphere shining by its own light, as was taught by Aristotle. Instead, he observed that it shone by reflected light, and was covered with hills, valleys, and craters. He also turned his telescope towards the Milky Way and discovered that the great band of faint light was in fact given by a great mass of stars. He said that the number of stars to be found there was "so numerous as to be almost beyond belief".

In 1610, Galileo discovered the four bright satellites (moons) of Jupiter, and shortly afterwards discovered that Venus goes through similar phases as the Moon. Mars was also observed to go through slight phases as well. He also discovered the existence of sunspots on the surface of the Sun, and observed the rings of Saturn (though he did not realize them for what they were).

Unfortunately for Galileo, his discoveries contradicted the views held by the people of the time. He was forced to renounce the Copernican theory of the Universe, and was sentenced to house arrest for the remainder of his life.

During the 1660's, Isaac Newton made a very important discovery - the law of gravitation. According to this law, every body in the universe attracts every other body, with higher masses resulting in a stronger attraction. In addition, the attraction is inversely proportional to the square of the distances between the objects. This crucial law of physics is summarized by the following equation:

```
                   G * M1 * M2
               F = -----------
                       r*r
   where:
          F = the force of attraction between two objects,
              given in units of "Newtons"
          G = the universal constant of gravitation, 6.672e-11 m^3/(kg s^2)
          M1= mass of the first body, given in kilograms
          M2= mass of the second body, given in kilograms
          r = distance between the two bodies, given in metres.
```

This law turned out to be basic to our understanding of the motions of objects in the universe. It provides an explanation as to why objects fall to the ground when they are dropped, or why the Moon stays in orbit around the Earth.

Newton also discovered that sunlight is made of a combination of many different colors. By passing light through a prism, he found that he could extract these colors into a rainbow. This discovery was the basis of the spectrograph, an instrument which was capable of measuring the spectrae of stars and ultimately determining their composition. For more information, see Electromagnetic Radiation.

Since the beginning of history, people knew of the six brightest planets - Mercury, Venus, Earth, Mars, Jupiter, and Saturn. There was no reason to expect any others. But in 1781, William Herschel discovered a seventh planet, which became known as Uranus. Since it was twice as far as Saturn, astronomers observed Uranus carefully to see if Newton's laws of motions worked at distances so far from the Sun. By 1783, it became clear that the planet had wandered slightly from its predicted orbit. It took until the next century to explain why this happened.

See also A Short History of Astronomy (ancient - A.D. 999)
A Short History of Astronomy (1800 - present)

1.4 A Short History of Astronomy (1800 - present)

Towards the end of the 18th century, astronomers had begun to notice that Uranus did not move exactly as predicted by Newton's theory of gravitation. Several astronomers, including John Adams and Urbain Leverrier thought that this unexplained motion may result from the gravitational influences of another, unknown planet. In 1846, Leverrier calculated where this planet might be located, and sent his results to the German astronomer Johann Galle. Neptune was discovered in the predicted place. See Neptune - Discovery for more information.

However, closer examination indicated that the planets were still not following the exact orbits as predicted by Newton. In 1905, Percival Lowell computed where another planet might be to cause the discrepancies. Clyde Tombaugh used these calculations to target regions of the sky and succeeded in discovering Pluto in 1930. See Pluto - Discovery for more information.

No other major objects in our solar system have since been discovered, though several hypothetical objects have been considered from time to time.

1905 was a very important year for astronomy and physics. During this period, Albert Einstein published three landmark papers in the German scientific periodical "Annalen der Physik" (Annals of Physics). Each of these three papers went on to become the basis of an entirely new branch of physics. They dealt with the topics of:

1. Quantum mechanics and the photoelectric effect.
2. The Special Theory of Relativity.
3. Brownian motion

Einstein also published his General Theory of Relativity in 1915. These concepts (especially the Theory of Relativity) are all covered in detail in the section on Einstein. His unique way of looking at the universe resulted in the most significant advances since Newton's time. Generally speaking, Newtonian physics were still valid when dealing with slow moving objects in a low gravitational field. However, in more extreme conditions, Einstein's theories of the interaction of objects were required.

In 1931, Karl Jansky discovered radio waves reaching the Earth with their origin in the Milky Way. This was the first indication that we were receiving electromagnetic energy from stars and other objects in the universe at wavelengths other than that of visible light. His discovery was of great practical importance to astronomers who were now capable of studying the energy given off by stars over a much broader range of the electromagnetic spectrum.

In 1948, the Australian astronomers John G. Bolton and Gordon J. Stanley discovered the first star which gave off radio waves but no visible light. In the following years, radio signals were found to originate from a wide variety of objects in the universe, ranging from massive galaxies to planets within our own solar system.

Ten years later, humankind made its first venture into space with the Russian launch of the Sputnik satellite. This was the first of literally thousands of launches, and was an early indication of an exciting future.

In 1960, astronomers discovered very distant objects which emitted incredibly large amounts of energy at visible and radio wavelengths. These became known as quasars. By the end of the decade, brief pulses of light and radio waves occurring at extremely regular intervals were detected from several objects. At first, some astronomers thought that these signals might be signs of some extraterrestrial intelligence, but before long they realized that they had detected pulsars.

In 1969, American astronauts set foot upon the moon with the mission of Apollo 11. After thousands of years of human development, we had finally reached a level where we could begin to visit some of the objects that humankind had wondered about for so long.

Stephen Hawking came to the forefront of physics research in 1974 when he suggested that miniature black holes the size of a proton could have existed shortly after the Big Bang. Due to their incredible mass and extremely small size, both Einstein's Theory of Relativity and quantum mechanical effects would have to be considered. Due to quantum mechanics, he realized that subatomic particles could periodically travel faster than light for short periods of time, effectively escaping from the mini black holes. These black holes would slowly "evaporate" until they exhausted their energy supply and finally exploded. Hawking is considered one of the foremost experts on black holes.

See also A Short History of Astronomy (1000 - 1799)
Interplanetary Spacecraft

1.5 Strobel, Nick

Though most of the hypertext information accompanying this software was created exclusively for "The Digital Universe", several chapters dealing with a general introduction to astronomy were provided by Nick Strobel. They have been used to present an astronomy course at the University of Washington and have been edited for inclusion with the software.

1.6 Hagen, William F.

In 1992, William Hagen wrote descriptions of a wide variety of Earth-orbiting satellites. The producers of "The Digital Universe" would like to thank him for his efforts, and have included his descriptions in the online encyclopedia. They are provided almost verbatim, with hypertext links added and minor spelling corrections made.

1.7 General Introduction and Naked Eye Astronomy

The following chapter was written by Nick Strobel. It has been included with "The Digital Universe" with the kind permission of the author.

INTRODUCTION TO ASTRONOMY
AND
ASTRONOMY WITHOUT A TELESCOPE

1.8 Introduction: The scale of things

INTRODUCTION: THE SCALE OF THINGS

We can start by setting the stage for what we're up against when we try to understand the physical universe in the discipline we call astronomy or astrophysics. This chapter will cover everything we notice by taking a look at the sky without the use of a telescope.

In order to get a feel for the time scales involved in astronomy, we can use the analogy of the cosmic calendar in which every second in the calendar corresponds to 475 normal years (so 24 calendar days = 1 billion regular years). Assuming that the universe is 15 billion years old, we can squeeze that huge timescale to one cosmic calendar year. Here are some important dates relevant to us humans:

```
Jan.  1 - Origin of the Universe
May   1 - Origin of our galaxy
Sep.  9 - Origin of our solar system
Sep. 14 - Earth formation
Sep. 30 - Life on Earth
```

Nov. 25 - Sexual reproduction advent
Dec. 1 - Oxygen atmosphere
Dec. 17 - Cambrian explosion (600 million years ago when most complex organisms, such as fish, trilobites, etc. appear)
Dec. 19 - Land plants and insects
Dec. 22 - First amphibians
Dec. 23 - First reptiles & trees
Dec. 25 - First dinosaurs
Dec. 30, 10:00 AM - KT impact, mammal age, birds
Dec. 30 - First primates
Dec. 31, 10:00 PM - Australopithicenes (Lucy, etc.)
Dec. 31, 11:25 PM - Homo habilis
Dec. 31, 11:40 PM - Homo erectus
Dec. 31, 11:50 PM - Early Homo sapiens
Dec. 31, 11:57 PM - Neanderthal man
Dec. 31, 11:58:38 PM - Cro-Magnon man
Dec. 31, 11:58:57 PM - Homo sapiens sapiens
Dec. 31, 11:59:39 PM - Human history began
Dec. 31, 11:59:55 PM - Ancient Greece

It is rather surprising that we've been able to discover so much about the long term evolution of the universe and the things in it, especially when you consider that we've only been seriously observing the universe for about 100 years, which is only a very slight fraction of the universe's lifetime. About 100 years ago is when photography was first used in astronomy, making truly systematic observation programs possible. How can we say that the Sun will go through a red giant phase in about 5 billion years from now with confidence? Is it hubris to confidently talk about the Earth's formation process about 4.5 billion years ago? To give an idea of the difficulties in studying long timespans consider this analogy: An alien comes to Earth to search for life and to understand how it evolved. ET has a camera and has just 15 seconds to take as many photographs as possible. Fifteen seconds is the same proportion of a human lifetime as the 100 years is to the universe's age (15 seconds/human lifetime = 100 years/universe age). ET returns home and its colleagues try to understand Earth from this 15 second period of snapshots. They won't see any important evolutionary changes. How will they determine the dominant life form? They could use a variety of criteria: 1) Size: leads them to choose whales or elephants; 2) Numbers: choose insects; 3) amount of land space controlled by one species: choose automobiles.

Suppose they somehow decide humans are dominant. They now have further problems. There is considerable diversity among the humans (though to an ET with 10 tentacles, 200 eyes, and a silicon outer shell, the humans all look alike!). ET and its colleagues try to systematically classify the humans. The humans come in a variety of sizes. In a coarse classification scheme, they break the sizes down into small, medium, and large. They also come in variety of optical colors for their outer shell: red, black, brown, yellow, and white. There appears to be 2 separate sexes. After some false starts with theories that used hair length and eye color, they are ready to ask themselves, "Do small, brown, female humans evolve into large, red, male humans?" "Do the small stay small and the large stay large?" "Why is there a tendency for small humans to be with one or two large humans?" With the three characteristics [size (3 divisions), color (5 divisions), and sex (2 divisions)], ET has 3x5x2 different combinations and 30x30 possible evolutionary schemes to consider! The universe has a lot more characteristics and, therefore, many more combinations to consider!

1.9 Science in General

SCIENCE IN GENERAL

The scientist usually learns about nature by using controlled experiments in which only one thing at a time is varied, and the experiments can be repeated by anyone as many times as they want to verify that the effect is reproducible. The astronomer cannot do controlled experiments. We cannot even examine things from a variety of angles. What we do is collect light and other radiation from celestial objects and use all of our information and creativity to interpret the signals from afar. We try to look for the experiments nature has set up for us and hone on a few basic characteristics at a time. Like any scientist, the astronomer makes observations, which suggest hypotheses. These speculations are made into predictions of what may be observed under slightly different observing and/or analysis circumstances. The astronomer returns to the telescope to see if the predictions pan out or if some revision needs to be made in the theory. Theory and observation play off each other.

Often the evidence for a particular hypothesis is indirect and will actually support other hypotheses as well. The goal is to make an observation that conclusively disproves one or more of the competing theories. Currently unresolvable questions may be resolved later with later improved observations using more sophisticated/accurate equipment. Sometimes new equipment shows that previously accepted theories and hypotheses are wrong!

1.10 Value of Astronomy in the Scientific Endeavor

VALUE OF ASTRONOMY IN THE SCIENTIFIC ENDEAVOR

However, not all is negative. The universe gives us a vast number of different phenomena to observe. Many of these things cannot be reproduced in Earth laboratories. There are gas clouds in such a rarefied state that they give off radiation not seen on Earth. Some objects are so dense that their gravitational fields bend light so much that it is prevented from leaving the object! Many things that are unlikely or impossible on Earth are routinely observed in the cosmos. Many of the scientific theories in other fields make predictions of what would happen under very extreme circumstances. Sometimes those extreme circumstances are the only situations distinguishing two or more contradictory theories. Unfortunately, the scientists of those other disciplines cannot test their "wild" ideas. Astronomy allows those theories to be tested. Very subtle but crucial processes may be missed by observers focusing on the Earth, but the astronomer can see those processes magnified to easily noticeable levels in some other celestial object.

Now back to the long term evolution side of the coin. We actually have a time machine! Not the H. G. Wells variety or G. Roddenberry’s Gate of Eternity but something much simpler due to the large distances and finite speed of light (300,000 km/sec). It takes time for radiation from a celestial object to reach us. Therefore, when we examine an object at a large distance from us, we see it as it was. The farther away the object is, the longer it took the radiation to reach us, and the further back in time we observe it. The Sun is 150 million kilometers from us, so we see the Sun as it was 8.3 minutes ago. The farthest object you can see without a telescope is the Andromeda galaxy about 1.9*10^19 km from us so we see it as it was a little over 2 million years ago.

Astronomres use a more convenient length scale than the kilometer. They use a light year which is equal to the distance light travels in one year (about 9.46*10^12 km). So the Andromeda galaxy is a little over 2 million light years away from us.

To study the evolution of long-lived objects like stars (with lifetimes of millions to billions of years) or galaxies, we observe the objects of interest at different distances from us so we see it at different epochs. We then see the object at various different ages or evolutionary stages.

1.11 Celestial sphere defined

CELESTIAL SPHERE DEFINED

Now that we have some feeling for the scales of time and space that astronomy encompasses and some of the difficulties caused by being Earth-bound, we can take a look at what is up there in the sky beyond the clouds. At first, we'll use a simple model of the sky that any observer would consider reasonable or at least workable. Imagine the sky as a great, hollow, sphere surrounding the Earth. The stars are attached to this sphere-some bigger and brighter than others-which rotates around the stationary Earth roughly every 24 hours. Alternatively, you can imagine the stars as holes in the sphere with light from the heavens beyond the sphere shining through those holes. Of course, we now know that a star's apparent brightness is determined by its distance as well as its physical size and temperature. We also know that it is actually the Earth that is rotating. This imaginary sphere is called the celestial sphere, and has a very large radius so that no part of the Earth is significantly closer to a given star than any other part. Therefore, the sky always looks like a great sphere centered on our position. The celestial sphere (and, therefore, the stars) appears to move westward-stars rise in the east and set in the west.

Why a sphere? The Earth is spherical! This was known much earlier than Columbus' time. Sailors had long known that as a ship sailed away from the shore it not only diminished in apparent size, but it also appeared to sink into the water. The simplest explanation to use was that the Earth was curved (particularly, since those ships did come back without falling off some edge!). They also knew that if one traveled in a north-south direction, some stars disappeared from view while others appeared. The simplest explanation said that the Earth is round, not flat. Pythagoras noted that the shadow of the Earth falling on the Moon during a lunar eclipse was always curved and the amount of the curvature was always the same. The only object that always casts a circular shadow regardless of its orientation is a sphere. We know about this Pythagorean argument through the writings of Aristotle.

1.12 Reference Markers

REFERENCE MARKERS

Now for some reference makers: The stars rotate around the North and South Celestial Poles. These are the points in the sky directly above the geographic north and south pole, respectively. The Earth's axis of rotation intersects the celestial sphere at the celestial poles. The number of degrees the celestial pole is above the horizon is equal to the latitude of the observer. Fortunately for observers in the northern hemisphere, there is a fairly bright star very close to the North Celestial Pole (Polaris or the North star). Another important reference marker is the celestial equator: an imaginary line around the sky directly above the Earth’s equator. It is always 90 degrees from the poles. All the stars appear to rotate in a path that is parallel to the celestial equator. The celestial equator intercepts the horizon at the points directly east and west anywhere on the Earth. From the North Pole, you would see Polaris straight overhead and the celestial equator on your horizon. The point directly above an observer is called the zenith and is always 90 degrees from the horizon.

For each degree you move south, the North Celestial Pole (NCP from here on) moves 1 degree away from the zenith toward the north and the highest point of the celestial equator’s curved path in the sky moves up one degree from the southern horizon. By the time you reach Edmonton, Canada (at latitude 53.5 degrees N) the NCP has moved 36.5 degrees away from the zenith so it is now 90 - 36.5 = 53.5 degrees above the horizon. The celestial equator is 36.5 degrees above the southern horizon and it still intercepts the horizon due east and west. At the equator, you would see the celestial equator arcing from east to the zenith to the west. The NCP would be on your northern horizon. Continuing southward, you would see the NCP disappear below the horizon and the SCP rise above the southern horizon one degree for every degree of latitude south of the equator you went. The arc of the celestial equator would have moved to the north, but the arc still intercepts the horizon at the east/west points.

1.13 Angles

ANGLES

To describe distances across the sky, we use "angles on the sky" as if the celestial objects were actually attached somehow to the celestial sphere. There are 360 degrees in a full circle and 90 degrees in a right triangle. Each degree is divided into 60 minutes of arc. A quarter viewed face-on from across the length of a football field is about 1 arc minute across. Each minute of arc is divided into 60 seconds of arc. The ball in the tip of a ballpoint pen viewed from across the length of a football field is about 1 arc second across. The Sun and Moon are both about 0.5 degrees = 30 arc minutes in diameter. The pointer stars in the bowl of the Big Dipper are about 5 degrees apart and the bowl of the Big Dipper is about 30 degrees from the NCP.

1.14 Motion of the Sun

MOTION OF THE SUN

Every day the Sun rises in an easterly direction and sets in a westerly direction, and takes on average 24 hours to go from noon position to noon position the next day. The "noon position" is when the Sun is highest above the horizon on a given day. Our clocks are based on this solar day. However, the exact position on the horizon of the rising and setting Sun varies throughout the year. Furthermore, the Sun appears to drift eastward with respect to the stars (or lag behind the stars) over a year’s time. It makes one full circuit of 360 degrees in 365.25 days (very close to 1 degrees or twice its diameter per day). The apparent path of the Sun through the stars is called the ecliptic. This circular path is tilted 23.5 degrees with respect to the celestial equator. The ecliptic and celestial equator intersect at two points: the vernal (spring) equinox and autumnal equinox. The Sun crosses the celestial equator moving northward at the vernal equinox around March 21 and crosses the celestial equator moving southward at the autumnal equinox around September 22. When the Sun is on the celestial equator at the equinoxes, everybody on the Earth experiences 12 hours of daylight and 12 hours of night (hence, the name "equinox" for equal night).

No matter where you are on the Earth, you will see 1/2 of the celestial equator's arc (except for the geographic poles). Since the sky appears to rotate around us in 24 hours, anything on the celestial equator takes 12 hours to go from due east to due west. Every celestial object's diurnal (daily) motion is parallel to the celestial equator. So for northern observers, anything south of the celestial equator takes less than 12 hours between rise and set, because most of its rotation arc around is hidden below the horizon. Anything north of the celestial equator takes more than 12 hours between rising and setting because most of its rotation arc is above the horizon. For observers in the southern hemisphere, the situation is reversed. However, remember, that everybody anywhere on the Earth sees 1/2 of the celestial equator, so at the equinox, when the Sun is on the equator, we see 1/2 of its rotation arc around us, and therefore we have 12 hours of daylight and 12 hours of darkness every place on the Earth.

The geographic poles are a special case. The celestial equator is right along the horizon and the full circle of the celestial equator is visible. Since a celestial object's diurnal path is parallel to the celestial equator, stars do not rise or set at the geographic poles. On the equinoxes the Sun moves along the horizon. At the North Pole the Sun "rises" on March 21st and "sets" on September 22. The situation is reversed for the South Pole.

Since the ecliptic is tilted 23.5 degrees with respect to the celestial equator, the Sun's maximum distance from the celestial equator is 23.5 degrees. This happens at the solstices. For observers in the northern hemisphere, the farthest northern point from the celestial equator is the summer solstice, and the farthest southern point is the winter solstice. The Sun reaches winter solstice around December 21 and we see the least part of its diurnal path all year-this is the shortest "day". The Sun reaches the summer solstice around June 21 and we see the greatest part of its diurnal path all year-this is the longest "day".

1.15 Solar and sidereal day

Solar and Sidereal Day

The fact that our clocks are based on the solar day and the Sun appears to drift eastward with respect to the stars (or lag behind the stars) by about 1 degrees per day means that if we look closely at the positions of the stars over a period of several days, we notice that according to our clocks, the stars rise and set 4 minutes earlier each day. Our clocks say that the day is 24 hours long, so the stars move around the Earth in 23 hours 56 minutes. This time period is called the sidereal day because it is measured with respect to the stars. One month later (30 days) a given star will rise 2 hours (30*4 minutes = 120 minutes) earlier than it did before. A year later that star will rise at the same time as it did today. Another way to look at it is that the Sun has made one full circuit of 360 degrees in a year of 365.25 days (very close to 1 degree per day). This means that, from noon to noon, the Earth has to turn nearly 361 degrees, not 360 degrees, in 24 hours. This makes the length of time for one rotation with respect to the background stars a little less than 24 hours on the clock.

See also: Time Standards.

1.16 Solar and Sidereal Time as Viewed from Space

Solar and Sidereal Time as Viewed from Space

Let’s jump to a more modern view and take a position off the Earth and see the Earth revolving around the Sun in 365.25 days and rotating on its axis every 23 hours 56 minutes. The Earth’s rotation plane is tilted by 23.5 degrees from its orbital plane which is projected against the background stars to form the ecliptic.

View picture

Note that the Earth’s rotation axis is always pointed toward the celestial poles. Currently the North Celestial Pole is very close to the star Polaris.

Imagine that at noon we have a huge arrow that is pointing at the Sun and a star directly in line behind the Sun. The observer on the Earth sees the Sun at its highest point above the horizon: it is on the arc going through the north-zenith-south points, which is called the meridian. The observer is also experiencing local noon. If the Sun were not there, the observer would also see the star on the meridian.

View picture

Now as time goes on the Earth moves in its orbit and it rotates from west to east (counterclockwise if viewed from above the north pole). One sidereal period later (23 hours 56 minutes) or one true rotation period later, the arrow is again pointing toward the star. The observer on the Earth sees the star on the meridian. But since the Earth has moved in its orbit, the arrow is not pointing at the Sun! In fact the Earth needs to rotate a little more to get the arrow lined up with the Sun. The observer on the Earth sees the Sun a little bit east of the meridian. Four minutes later or one degree of further rotation aligns the arrow and Sun and we have one solar day (24 hours) since the last time the Sun was on the meridian. That night the Earth observer will see certain stars visible like those in Taurus, for example. A half of a year later Taurus will not be visible but those stars in Scorpius will be visible.

1.17 Time Zones

Time Zones

People east of you will see the Sun on their meridian before you see it on yours. For each one degree in longitude East a person is from you, the time between their local noon and yours will increase by 4 minutes. It used to be that every town's clocks were set according to their local noon but this got very confusing for the railroad system A more sensible clock scheme based on time zones was adopted. Each person within a time zone has the same clock time. Each time zone is roughly 15 degrees wide which corresponds to 15*4 minutes = 60 minutes = 1 hour worth of time. Those in the next time zone east of you have clocks that are 1 hour ahead, other than a few political exceptions.

1.18 Equation of Time

Equation of Time

This subsection can be skipped for those in a hurry. The careful observer will notice some peculiarities with the solar motion and will want to read this paragraph.

There is a further complication in that the actual Sun's drift against the stars is not uniform. Part of the non-uniformity is due to the fact that in addition to the general eastward drift among the stars the Sun is moving along the ecliptic northward or southward with respect to the celestial equator. Thus, during some periods the Sun appears to move eastward faster than during others. Apparent solar time is based on the component of the Sun's motion parallel to the celestial equator. This effect alone would account for as much as 9 minutes difference between the actual Sun and a mean Sun moving uniformly along the celestial equator. Another effect to consider is that the Earth's orbit is elliptical so when the Earth is at its closest point to the Sun (perihelion) it moves quickest. When at its farthest point from the Sun (aphelion) it moves slowest. Remember that a solar day is the time between meridian passages of the Sun. At perihelion the Earth is moving rapidly so the Sun appears to move quicker eastward than at other times of the year. The Earth has to rotate through a greater angle to get the Sun back to local noon. This effect alone accounts for up to 10 min difference between the actual Sun and the mean Sun. However, the maximum and minimum of these two effects do not coincide so the combination of the two (called the Equation of time) is a bit complicated. This explains why the earliest sunset and latest sunrise is not exactly on the winter solstice. Yet, the shortest day is at the winter solstice. Rather than resetting our clocks every day to this variable Sun, our clocks are based on a uniformly moving Sun that moves at a rate along the celestial equator of 360 degrees/365.2564 per day.

1.19 Seasons

Seasons

The seasonal temperature depends on the amount of heat we receive from the Sun. To hold the temperature constant, there must be a balance between the amount of heat we gain and the amount we radiate to space. If we receive more heat than we lose, we get warmer; if we lose more than we gain, we get cooler. Consider some basic observations of the seasons:

1) The northern hemisphere gets warmer during June-August (northern

summer) while the southern hemisphere get cooler at the same time (southern winter).

2) Shadows tend to be longer and days tend to be shorter during the winter months.

3) The Suns seems to rise and set further north during the northern hemisphere's summer. If you're in the southern hemisphere, during your summer (December-February) the Sun rises and sets further south.

4) The Sun seems to appear higher in the sky at local noon during the summer.

View picture

1.20 Motions of the Moon

MOTIONS OF THE MOON

The Moon moves rapidly with respect to the background stars. It moves about 13 degrees (26 times its apparent diameter) in 24 hours (slightly more than its own diameter in one hour)! Its rapid motion has given it a unique role in the history of astronomy. For thousands of years it has been used as the basis of calendars. Isaac Newton got crucial information from the Moon's motion around the Earth for his law of gravity. Almost everyone has noted that we see the same face of the Moon all of the time. It is often referred to as the "man in the Moon", "woman in the Moon", "rabbit in the Moon" etc. One thing this shows us is that the Moon turns exactly once on its axis each time that it goes around the Earth. Later on we'll find out how tidal forces have caused this face-to-face dance of the Earth and Moon. It drifts eastward with respect to the background stars (or it lags behind the stars), and returns to the same position with respect to the background stars every 27.323 days. This is its sidereal period.

1.21 Phases and Eclipses

Phases and Eclipses

One of the most familiar things about the Moon is that it goes through phases from new (all shadow) to first quarter (1/2 appears to be in shadow) to full (all lit up) to third quarter (opposite to the first quarter) and back to new. This cycle takes about 29.53 days. This time period is known as the Moon's synodic period. Because the Moon moves through its phases in about four weeks, new Moon, first quarter, full Moon, third quarter, and new Moon occur at nearly one-week intervals. We know that the phases are due to how the Sun illuminates the Moon and the relative positioning of the Earth, Moon, and Sun. We observe that not much of the Moon is illuminated when it is close to the Sun. In fact, the smaller the angular distance between the Moon and the Sun, the less we see illuminated. When the angle is within about 6 degrees we see it in a new phase. Sometimes that angle = 0 degrees and we have a solar eclipse-the Moon is in new phase and it is covering up the Sun. Conversely, the greater the angular distance is between the Moon and the Sun, the more we see illuminated. Around

180 degrees we see the Moon in full phase. Sometimes (about twice a year) the Moon-Sun angle is exactly 180 degrees and we see the Earth’s shadow covering the Moon-a lunar eclipse.

Why are the synodic and sidereal periods not equal to each other? The reason is similar to why the solar day and sidereal day are not the same. Remember that a solar day is slightly longer than a sidereal day because of the Sun’s apparent motion around the Earth (caused by the Earth’s motion around the Sun). The Moon’s synodic period is longer than its sidereal period because of its motion around the Earth. At new Moon, the Sun and Moon are seen from the Earth against the same background stars. One sidereal period later, the Moon has returned to the same place in its orbit and to the same place among the stars, but the in the meantime, the Sun has been moving eastward, so the Moon has not yet caught up to the Sun. The Moon must travel a little over two more days to reach the Sun and establish the new Moon geometry again.

View picture

The modern model has the Moon going around the Earth with the Sun far away. At different positions in its orbit we see different phases all depending on the relative positions of the Earth, Moon, and Sun.

1.22 Relationship of Tides and Lunar Phases

Relationship of Tides and Lunar Phases

When you look in the paper at the section containing the tide tables, you’ll often see the phase of the Moon indicated as well. That’s because the ocean tides are caused by the gravity of the Moon. This document will go into the details of gravity in more depth later, but an important concept at this point is that gravity depends on mass and distance--greater distance means less gravity. The side of the Earth facing the Moon is about 4000 miles closer to the Moon than the center of the Earth is, and the Moon’s gravity pulls on the near side of the Earth more strongly than on the Earth’s center. This produces a tidal bulge on the side of the Earth facing the Moon. The Earth rock is not perfectly rigid; the side facing the Moon responds by rising toward the Moon by a few centimeters on the near side. The more fluid seawater responds by flowing into a bulge on the side of the Earth facing the Moon. That bulge is the high tide. At the same time the Moon exerts an attractive force on the Earth’s center that is stronger than that exerted on the side away from the Moon. The Moon pulls the Earth away from the oceans on the far side, which flow into a bulge on the far side, producing a second high tide on the far side.

View picture

These tidal bulges are always along the Earth-Moon line and the Earth rotates beneath the tidal bulge. When the part of the Earth where you are located sweeps under the bulges, you will notice a high tide; when it passes under one of the depressions, you experience a low tide. An ideal coast should experience the rise and fall of the tides twice a day. In reality, the tidal cycle also depends on the latitude of the site, the shape of the shore, winds, etc.

The Sun’s gravity also produces tides that are about half as strong as the Moon’s and produces its own pair of tidal bulges. They combine with the lunar tides. At new and full Moon, the Sun and Moon produce tidal bulges that add together to produce extreme tides. These are called spring tides. When the Moon and Sun are at right angles to each other (1st & 3rd quarter), the solar tides reduce the lunar tides and we have neap tides.

1.23 Tides Slow Earth's Rotation

Tides Slow Earth's Rotation

As the Earth rotates beneath the tidal bulges, it attempts to drag the bulges along with it. A large amount of friction is produced which slows down the Earth's spin. The day has been getting longer and longer by about 0.0016 seconds each century. Over the course of time this effect can have a noticeable effect. Astronomers trying to compare ancient solar eclipse records with their predictions found that they were off by a significant amount. But when they took the slowing down of the Earth's rotation into account, their predictions agreed with the solar eclipse records. Also growth rings in ancient corals about 400 hundred million years old show that the day was only 22 hours long so that there were over 400 days in a year. Eventually the Earth's rotation will slow down to the point where it keeps only one face toward the Moon. Gravity acts both ways so the Earth has been creating tidal bulges on the Moon and has slowed it's rotation down so much that it rotates once every orbital period. The Moon keeps one face always toward the Earth.

1.24 Tides Enlarge the Moon's Orbit

Tides Enlarge the Moon's Orbit

Friction with the ocean beds drags the tidal bulges eastward out of a direct Earth-Moon line and since these bulges contain a lot of mass, their gravity pulls the Moon forward in its orbit. The Moon's orbit is growing larger, receding from the Earth at about 3 cm per year. We have been able to measure this slow spiraling out with lasers bouncing off reflectors left by the Apollo astronauts on the lunar surface. The consequence of the Moon's recession from the Earth because of the slowing down of the Earth's rotation is also an example of the conservation of angular momentum. Angular momentum is the amount of spin motion an object or group of objects has. It depends on the geometric size of the object, how fast the object is moving, and the mass of the object. The slow spiraling out of the Moon means that there will come a time in the future when the angular size of the Moon will be smaller than the Sun's and we won't have any more solar eclipses!

1.25 Lunar Eclipses

Eclipse Details: Lunar Eclipse

Let's talk a little more about lunar and solar eclipses. Remember that an eclipse happens when an object passes through another object's shadow. Any shadow consists of two parts: an umbra which is the region of total shadow and the penumbra which is the outer region of partial shadow. If the Moon were to pass through the Earth's umbra, a Moon observer would not be able to see the

Sun at all-she would observe a solar eclipse! An Earth observer would see a total lunar eclipse. The Earth’s shadow is pretty big compared to the Moon so a total lunar eclipse lasts about 1 hour and 45 minutes. If the Moon only passed through the outer part of the shadow (the penumbra) then the Moon observer would see the Sun only partially covered up-a partial solar eclipse. The Earth observer would see the Moon only partially dimmed-a partial lunar eclipse. During a total lunar eclipse we see another interesting effect-the Moon turns a coppery (or bloody) red. This is due to sunlight refracting or bending through the Earth’s atmosphere. Dust particles in the Earth’s atmosphere have removed much of the bluer colors in the sunlight so only the redder colors make it to the Moon. The amount of dust determines the deepness of the red colors. This is also why the Sun appear redder at sunset on Earth. The Moon observer would see a reddish ring around the Earth during a lunar eclipse.

1.26 Solar Eclipses

Eclipse Details: Solar Eclipse

The Moon’s shadow also has an umbra and penumbra. The shadow is much smaller than the Earth’s. Only if the Moon is in the ecliptic plane when it is exactly New Moon will we have the Moon’s shadow hitting the Earth. Where the umbra hits the Earth, we’ll see a total solar eclipse. Where the penumbra hits the Earth, we’ll see a partial solar eclipse. In a total solar eclipse the bright disk of the Sun is completely covered up by the Moon and we see the other parts of the Sun like the corona, chromosphere, and prominences. Unfortunately, only the tip of the Moon’s umbra reaches the Earth (the tip hitting the Earth is typically only 270 km [168 miles] in diameter) and it zips along the Earth’s surface at over 1600 kph (1000 mph) as the Moon moves around the rotating Earth. As a result, a total solar eclipse can last a maximum of only 7.5 min. Usually total solar eclipses only last 3-4 minutes. Because of the orbital motion of the Moon and the rotation of the Earth, the umbra makes a long, narrow path of totality.

Sometimes the umbra does not reach the Earth at all (only the penumbra does) even though the Moon is on the ecliptic and it is exactly in New Moon phase. This occurs if the Moon is too far to completely obscure the Sun’s disk. We see a bright ring around the Moon when it is lined up with the Sun-an annular eclipse (because of the annulus of light around the Moon).

1.27 Planetary Motions

PLANETARY MOTIONS

There are other celestial objects that drift eastward with respect to the stars. They are the planets (Greek for "wanderers"). There is much to be learned from observing the planetary motions with the naked eye (no telescope). There are 5 planets visible without a telescope: Mercury, Venus, Mars, Jupiter, and Saturn (6 if you include Uranus for those with sharp eyes!). All of them move within 7 degrees of the ecliptic. Two of the planets (Mercury and Venus) are never far from the Sun. Venus can get about 48 degrees away from the Sun, while Mercury can only manage a 27.5 degrees separation from the Sun. When Venus and/or Mercury are east of the Sun, we’ll see them as an "evening star" even though they are not stars at all. When either of them is west of the Sun they are called a "morning star". Planets produce no visible light of their own; we see them by reflected sunlight. True stars produce their own visible light. Venus can be the brightest of all the planets, sometimes getting so bright that it can create a shadow! Mercury and Venus are never visible at around midnight (or opposite the Sun), but the other planets can be visible then. Sometimes a strange thing happens-a planet will slow down its eastward drift among the stars, halt, and then back up and head westward for a few weeks or months (retrograde motion), then halt and move eastward again. The planet would execute a loop against the stars! This is caused by the geometry involved in viewing the moving planets from our moving Earth.

1.28 Time Standards

To establish a system of time measurement, two quantities need to be defined. The duration of a time interval (ie. the second) must be specified as well as the epoch, or the starting point, of the time. In astronomy, there are four types of systems in use, each with its advantages and disadvantages. These systems are:

```
Atomic Time    (TAI)
Dynamical Time (TD, TDT, TDB)
Sidereal Time  (GST, LST)
Universal Time (UT, UT0, UT1, UTC, LT)
```

1.29 Atomic Time (TAI)

With atomic time, the duration of a time interval corresponds to a defined number of wavelengths of radiation of a specific atomic transition of a chosen isotope. This fundamental unit is currently defined as the Système International (SI) second, and corresponds to the duration of 9,192,631,770 periods of the radiation corresponding to a transition between two hyperfine levels of the ground state of cesium-133.

International Atomic Time (Temps Atomique International, or TAI) is determined from a variety of atomic clock references around the world. First, an intermediate time scale (Échelle Atomique Libre, or EAL) is formed from a weighted average of the different clocks. Then, corrections are applied for known effects to maintain a unit of time as close as possible to the definition of the SI second. The result is TAI.

The SI second and TAI are often used as bases for the interpolation and prediction of other timescales.

See also: Time Standards.

1.30 Dynamical Time (TD)

Dynamical Time has a unit of duration based upon the orbital motion of objects within our solar system. It is closely related to Atomic Time. At the IAU General Assembly of 1976 in Grenoble, France, the following points were agreed

```
upon for a definition of Terrestrial Dynamical Time (TDT):

 1.  On 1977 January 1 00:00:00 TAI, the value of the new timescale for
     apparent geocentric ephemerides will be 1977 January 1 00:00:32.184
     exactly.
 2.  The unit of this timescale will be a day of 86,400 SI seconds at
     mean sea level.  According to Einstein's Theory of Relativity,
     the rate of progression of time depends upon the intensity of the
     gravitational field an object is in.  As a result, the altitude
     of a time source above the Earth's surface is important in the definition.
Two different types of Dynamical Time are defined:


 Terrestrial Dynamical Time (TDT)
 --------------------------------
 This time is that defined above for geocentric observations by the IAU. Since
its unit of timescale is defined in essentially the same manner as TAI, it can
roughly be defined as a simple offset from TAI as follows:

              TDT = TAI + 32.184 seconds
Note, however, that TAI is based on a statistical sampling of a large number of
atomic clocks, whereas TDT is an idealized uniform timescale.  As a result, the
unit of timescale is not perfectly identical, though the difference is not
significant for most practical purposes.


 Barycentric Dynamical Time (TDB)
 --------------------------------
 Einstein's Theory of Relativity states that the rate of progression of time,
from the point of view of an outside observer, is dependent upon the
gravitational potential in which the source is found, as well as the source's
relative velocity with respect to the observer.  As a result, the dynamical
time as reckoned from the barycenter (center of mass) of the solar system is
slightly different than that reckoned from the Earth (upon which TDT is
defined).  TDB may be approximated by:

              TDB = TDT + 0.001658*sin(g)+0.000014*sin(2*g)
   where:
              g=357.53+0.9856003*(JD-2451545.0) degrees
              TDB, TDT are expressed in seconds.
As can be seen, the difference between TDB and TDT is 0.0017 seconds at most.
Thus, for most practical observations, the distinction need not be made, and
simply Dynamical Time (TD) is quoted.

Dynamical Time has only been in use since 1984.  Before then, ephemeris time
was used in calculating ephemerides.

See also: Time Standards.
```

1.31 Sidereal Time

```
Though the Sun appears to move around the Earth roughly once every 24 hours,
the combination of the Earth's rotation around its axis and revolution around
the Sun make the stars appear to take only 23 hours and 56 minutes to do so.
As a result, the day reckoned by the stars, or the sidereal day, is shorter
```

than the solar day of 86,400 SI seconds. More precisely, the sidereal day is about 0.9972697 times as long as the solar day.

The Earth's rotation is gradually slowing, and doing so irregularly. As a result, the progression of sidereal time is not completely uniform when compared against TAI.

Two "types" of sidereal time are in common use:

```
 Universal Sidereal Time (UST)
 -----------------------------
 This is the sidereal time in Greenwich, England (0 degrees longitude).  Noon
on a given sidereal day is that moment when the 0 hour line of right ascension
transits, or appears directly over an observer located there.


 Local Sidereal Time (LST)
 -------------------------
 This is defined in a similar manner to Universal Sidereal Time, except that
sidereal noon is that moment when the 0 hour line of right ascension passes
above an observer located at any point on the Earth's surface.  LST and GST may
be simply related by the following formula:

             LST = GST + east longitude/15.0
   where:
             LST,GST are measured in hours
             east longitude is measured in degrees
```

Since a sidereal day is shorter than a solar day, a unique moment in time can not be specified by a sidereal time and solar day. As a result, the concept of a sidereal date has been advanced. The Universal Sidereal Date is defined as the interval in sidereal days that has elapsed on the Greenwich meridian since the beginning of the sidereal day that was in progress at JD 0.0. "The Digital Universe" uses this concept of Universal Sidereal Date in its conversions to and from sidereal time.

See also: Solar and sidereal day, Time Standards.

1.32 Universal Time (UT)

Universal Time forms the basis for all civil timekeeping, and conforms closely to the motion of the Sun. However, since the Earth is gradually slowing down its rate of rotation, and doing so erratically, it differs from TDT by the factor:

```
            Delta T = TDT - UT
```

This difference cannot be predicted by mathematical formulae, but must be determined by observation.

From observations of the diurnal motions of the stars, astronomers can determine the current universal time, uncorrected for the location on the Earth at which it was measured. This time is denoted UT0. The exact latitude and longitude of an observer depends upon the position of the poles of the Earth. It has been observed that the Earth's axis is not fixed in place, but moves around the axis of maximum moment of inertia in a phenomenon known as polar

```
motion.  It cannot be predicted accurately, but causes the position of the
poles to shift by about 9 metres on the surface of the Earth over a period of
roughly six years.

When corrections are made to UT0 for the effects of polar motion upon the
observations, the resulting timescale is called UT1.  It is influenced by the
slightly variable rotation of the Earth.

UT1 can be defined in terms of the Universal Mean Sidereal Time as follows:

 UMST of 0 hours UT1 = 24110.54841+8640184.812866*T+0.093104*T^2-6.2e-6*T^3
 where:
       T=d/36525, d being the number of days of Universal Time elapsed
                  since JD 2451545.0 and taking on values of
                  +-0.5, +-1.5, +-2.5, etc.
       UMST is expressed in seconds.
 Coordinated Universal Time (UTC)
 --------------------------------
 Since January 1, 1972, most broadcast signals began to distribute the time
based on UTC.  This timescale differs from TAI by an integer number of seconds,
and is kept to within 0.90 seconds of UT1 by the introduction of leap seconds
when required to compensate for the Earth's erratic rotation.  These are
usually added or removed at the end of June or December.


 Local Time (LT)
 ---------------
 Since most of the world does not live in Greenwich, a standard time is
required which is practical for other locations on Earth.  At one time, each
community had its own local time, defined so that noon occurred when the Sun
was roughly overhead.  But when railroads were built, a mass of confusion arose
in the resulting timetables.  To overcome the difficulty, scientists conceived
of a notion where the world would be divided into 24 "time zones", each
spanning 15 degrees in longitude.  Each time zone would be offset by an integer
number of hours from UTC. All cities within a given time zone would keep their
clocks set to the same time, and noon would occur within an hour of when the
Sun was overhead.

In practice, political boundaries caused the time zone boundaries to be drawn
rather irregularly instead of uniformly along lines of longitude.  Some regions
decided to keep their clocks set to a non-integral number of hours from UTC
(being offset by an additional 15 or 30 minutes).  Daylight Savings Time was
introduced in many (but not all) areas where clocks were moved an hour (or two)
ahead of their normal settings in the summer.  Imperfect as it may be, it is
the situation in which we currently find ourselves.
```

1.33 Electromagnetic Radiation (light)

```
The following chapter is based on information provided by Nick Strobel. It has
been edited, and included with "The Digital Universe" with the kind permission
of the author.

                  NOTES ON LIGHT (ELECTROMAGNETIC RADIATION)

     * General Properties: Frequency, spectrum, temperature
```

* Production of Light
    + Continuous Radiation
    + Discrete Radiation Emission lines and absorption lines
* Bohr model of the atom
    + Bohr model and light production

1.34 General Properties: Frequency, spectrum, temperature

Light, electricity, and magnetism are manifestations of the same thing called electromagnetic radiation. General properties:

1. Light can travel through empty space

2. The speed of light is constant in space. All forms of light travel at the same speed in space; they are (from highest energy to lowest energy): gamma rays, X-rays, Ultraviolet, Visible, Infrared, Radio. (Microwaves are high energy radio waves.)

3. The wavelength of light is defined similarly to that of water waves-distance between crests or between troughs. Visible light (what your eye detects) has wavelengths of 4000-8000 Angstroms. 1 Angstrom = 10^-8 cm.

White light is made of different colors (wavelengths). Not all wavelengths of light from space make it to the surface of the Earth. Long wave UV, visible light, parts of IR, and parts of the radio spectrum make it to the surface. More IR reaches mountains above 9,000 feet (2765 m) elevation. Gamma rays, X-rays, most of UV, most of IR, and other parts of the radio spectrum are only visible from above the atmosphere.

Frequency--number of wave crests that pass by a point every second. For light we find the simple relation between the speed of light c, wavelength (l) , and frequency (f) :

f=c/l

Blue light - shorter l , higher f, more energy.

Red light - longer l , lower f, less energy.

Energy of light = h*f (h is Planck's constant) = h*c/l

Spectrum - display of the intensity of light (EM radiation) at different wavelengths or frequencies

Temperature - measure of the random motion (or energy) of matter. A higher temperature means more motion (or energy). A natural scale would have zero motion at zero degrees (absolute zero). This scale is the Kelvin scale. The scale is exactly like the Celsius system but offset by 273 degrees.

1.35 Production of Light

```
Production of Light

Objects can either give off electromagnetic radiation at all wavelengths or
just at specific, discrete wavelengths:

  CONTINUOUS RADIATION
     * Continuous Radiation - also called Blackbody radiation or
       Thermal radiation. Produced by solids, liquids, and dense
       (thick) gases. Some important features:
         1. Light is generated from an object at all possible wavelengths
            if the object is above 0 K (absolute zero).
         2. The Shape of the continuous spectrum only depends on the
            temperature of the object and not on its chemical composition.
         3. A higher temperature means more light at all wavelengths.
         4. A higher temperature means that the peak of the curve is shifted
            to smaller wavelengths (higher frequencies). [l*T = 0.29 cm K.]
         5. The amount of energy passing through a square centimeter every
            second = sigma*T^4 , where sigma is a constant of nature.
            Note the fourth power! A small rise in temperature means a
            large rise in the luminosity (amount of energy/sec).


  DISCRETE RADIATION
     * Discrete Radiation - Emission and Absorption lines. Produced by
       low density (thin) gases. Some important features:
         1. Spectral features (either emission or absorption lines) are
            at certain discrete wavelengths (colors) and nowhere else.
         2. The wavelengths emitted depend on chemical composition which
            means that each element or molecule has a distinct pattern of
            lines.  We can get a "fingerprint" of the elements or molecules
            in the thin gas.


       Emission lines - light is produced.

       Absorption lines - light is taken out (absorbed) from a
       continuous source.  Absorption lines need a continuous source in
       the background, whereas emission lines do not.

       A spectrum from an object can contain continuous radiation,
       emission lines, and absorption lines simultaneously!

       If there is a thin gas between you and a continuous source, what
       type of discrete features you see depends on the relative temperature
       of the thin gas and continuous source. If the temperature of the thin
       gas is greater than the temperature of the continuous source,
       we see a continuous spectrum with emission lines. If the converse
       is true, we see a continuous spectrum with absorption lines.
```

1.36 The Bohr model of the atom

```
The Bohr model of the atom
```

The so-called "Bohr" model of the atom is very useful in explaining how emission and absorption lines are found in spectrae.

In this model, the central nucleus of the atom is made of protons (positive charge) and neutrons (neutral or zero charge) with electrons in certain specific orbits (energies) around nucleus. In a neutral atom, the number of electrons is equal to the number of protons.

An ion has an extra positive or negative charge. For example, the carbon ion C+ has 6 protons and 5 electrons, leaving it with a net positive charge.

An atom must follow certain "rules":

1. An electron can occur only at certain energies corresponding to distances from the nucleus. It is like an energy ladder, in which the electrons can lie only at rungs of the ladder and not in between rungs.
2. The orbits closest to the nucleus have the lowest energy. The lowest orbit is called the ground state.
3. Atoms want to be in the lowest energy state (electrons as close to the nucleus as possible).

How emission lines are created:

An atom starts out in an "excited" state. In this state the electron is not in the lowest energy orbit possible. From Rule # 3 above, it will attempt to much to the ground state, but Rule # 2 indicates that it will lose a specific amount of energy in the process. The energy of the photon released is equal to the energy difference between the two orbits.

View picture

How absorption lines are created:

An atom is hit by a photon of a certain specific energy (frequency). The atom absorbs the energy and pops an electron to a higher orbit. From Rule # 3, the photon will later drop back to its original orbit, releasing the original amount of energy absorbed. However, the photon will be released in a random direction - not necessarily in the same direction as the original photon. As a result, the observer sees less photons from the continuous source at that specific frequency (color) than at other frequencies. The spectrum appears continuous except for dark bands where the energy had been absorbed.

View picture

1.37 Telescopes

The following chapter is based on information provided by Nick Strobel. It has been edited, and included with "The Digital Universe" with the kind permission of the author.

TELESCOPES

```
    * Types of telescopes: Definition of telescope
        1. Refractor
        2. Reflector
    * Powers of a telecope
        1. Light Gathering Power
        2. Resolving Power
        3. Magnification
    * Atmospheric distortion
        1. Seeing
        2. Reddening and Extinction
        3. Atmospheric lines
```

1.38 Types of Telescopes

```
Types of Telescopes

Telescopes are instruments used to gather and focus light. Visible light and
large parts of the radio spectrum can be observed anywhere on the ground, but
small parts of the infrared (IR) must be observed above the water layer in the
atmosphere (at elevations above 2750m, or 9000 ft). Gamma rays, X-rays, UV,
most IR, and the rest of the radio frequencies must be observed in space.

The part of the telescope collecting the light is called the objective.

There are two main classes of telescopes:

1.  REFRACTING TELESCOPES

These use a lens to bend light.  They were the first ones built.
  Disadvantages:
    a.  Chromatic Aberration - redder colors are bent less than blue colors,
        so we see a rainbow of colors around the image.  Multiple compensating
        lenses or extremely long objective focal lengths are required to
        minimize the effect.  This is why the early refracting telescopes
        were made very long.
        View picture
    b.  The support for the lens is only around the edges of the lens.  As
        a result, it tends to sag under its own weight.  A lens 100
        centimetres in diameter is the largest refracting objective built.


2.  REFLECTING TELESCOPES

These use a mirror to reflect the light.  All modern research telescopes are of
this type.  They have a parabolic-shaped mirror which focuses parallel light
rays to a single point (remember that celestial objects are so far away that
all the light rays from an object hit the Earth as parallel rays). The focus is
before the eyepiece, so the image in astronomical telescopes (even with the
refractor above) is upside down.  Terrestrial viewing telescopes use other
lenses to re-invert the image right side up.
  Advantages:
    a.  No chromatic aberration
    b.  Support for the object is along the entire back of the mirror, so the
        mirrors can be made extremely large (up to 5 metres in diameter).
  Disadvantages:
```

a. Can be the victims of "spherical aberration". If the mirror is not curved enough, all of the light is not focused to a single point. The Hubble Space Telescope objective suffers from this problem, so it uses corrective optics to compensate.

View picture

1.39 Powers of a telescope

Powers of a telecope

Telescopes can be rated by several different "powers", some of which are more significant than others:

1. LIGHT GATHERING POWER

A telescope acts as a light bucket, so the bigger the objective, the more light is collected and the image is brighter. Faint objects are only seen with big objective telescopes.

2. RESOLVING POWER

The resolving power is the ability of a telescope to see really small details, so objects that are close together in the sky are easily seen as separate. The theoretical minimum resolvable angle in degrees can be given by:

```
              180 * observation wavelength
                angle = -----------------------------
                        Pi * objective diameter
```

where the observation wavelength and objective diameter are given in the same units of length. Radio wavelengths are large, so radio telescope objectives must also be large in order to get decent resolving power. The atmosphere will usually smear images so this theoretical resolving power will not be reached with ground-based telescopes. Techniques such as "speckle interferometry" or "adaptive optics" can be used to help remove the effects of atmospheric distortion.

3. MAGNIFICATION

Magnification is the ability of telescopes to make an image bigger. It is the least important "power" because it magnifies telescope and atmospheric distortions so that a small, fuzzy, faint blob only becomes a big, fuzzy, fainter blob. The magnifying power of a telescope can be given by:

```
                         focal length of objective
         Magnification = -------------------------
                         focal length of eyepiece
```

where both focal lengths are provided in the same units.

1.40 Atmospheric Distortion

Atmospheric Distortion

There are a number of terms and concepts that relate to how the image from a telescope is affected from atmospheric distortion:

1. SEEING

The air is in turbulent motion and light from celestial objects is bent randomly in many ways over time periods of tens of milliseconds. Images dance about (twinkle) and appear blurred. This atmospheric effect is called "seeing". Good seeing is when the air is stable and the twinkling is small.

2. REDDENING AND EXTINCTION

The air also absorbs and scatters different wavelengths in different amounts. Redder light is scattered less by atmosphere molecules and dust than bluer light. Since blue light is scattered more, we have a blue sky and objects appear reddened when viewed through lots of air. Some wavelengths of light are scattered a lot and do not make it to the surface at all. Astronomers put their telescopes on high mountaintops to look through less atmosphere and see less distortion. But even then, the atmosphere degrades the image somewhat. The best observations are done from space, as is done with the Hubble Space Telescope.

3. ATMOSPHERIC LINES

Gases in the atmosphere can create absorption lines which appear to originate from a celestial object. Astronomers need to remove these atmospheric spectral lines from spectroscopy data, or they would be determining the composition of the Earth's atmosphere instead of the object they wish to study.

1.41 Determining the Properties of the Stars

The following chapter is based on information provided by Nick Strobel. It has been edited, and included with "The Digital Universe" with the kind permission of the author.

DETERMINING THE PROPERTIES OF THE STARS

* Stars--What Are They Like?
    + Distances
    + Inverse Square Law of Light Brightness
    + Magnitude System
    + Star Colors (Temperature)
    + Stellar Composition
    + Velocity of Stars: Doppler Effect
    + Stellar Masses
    + Stellar Radius
* Types of Stars and HR diagram
    + Temperature dependence
    + Spectral Types
    + Hertzsprung-Russell diagram Color-Magnitude diagram
* Spectroscopic Parallax
    + Spectral Type to Distance

```
          + Cluster Main Sequence
```

1.42 Distances

DISTANCES

We use trigonometric parallax to measure the distances of the nearby stars. If you look at a nearby object from different vantage points, it appears to shift against more distant background objects. The farther apart the vantage points are, the greater the shift. The farther away the object is, the less it appears to shift. Since the shifts are so small, we use arc seconds as the unit of their angular shift. The distances are usually quoted in parsecs (abbreviated with "pc"). One parsec = 206,265 Astronomical Units (A.U.= mean distance between the Earth and the Sun or 1.496*10^8 km). It is also equal to 3.26 light years. We can measure shifts as small as 1/50 arc seconds or 50 parsecs in distance.

View picture

1.43 Inverse Square Law of Light Brightness

INVERSE SQUARE LAW OF LIGHT BRIGHTNESS

The Inverse Square Law of Light Brightness states that energy from any light source radiates out in a radial direction so that concentric spheres (centered on the light source) have the same amount of energy pass through them every second. As light moves outward, it spreads out to pass through each square centimeter of those spheres. The amount of energy passing through sphere-1 surface = amount of energy passing through sphere-2 surface. The surface area of a sphere = 4*pi*r^2 , where r is the radius of the sphere. The amount of energy passing through a surface = flux * surface area. So sphere-1 flux * 4*pi*r1^2 = sphere-2 flux * 4*pi*r2^2 which means flux-1 = flux-2 * (r2^2/r1^2) . This is known as the inverse square law! The flux is the amount of energy reaching each square centimeter of a detector (eg., your eye, CCD, piece of the sphere) every second.

1.44 Magnitude system

MAGNITUDE SYSTEM

We specify the brightness of stars with the magnitude system. The magnitude system is based on how the human eye perceives differences in brightnesses, so it is logarithmic. Fainter objects have larger, more positive magnitudes. The apparent brightness of a star observed from the Earth is called the apparent magnitude. The apparent magnitude is a measure of the star’s flux. If the star was at a distance of 10 parsecs from us, then its apparent magnitude would equal its absolute magnitude. The absolute magnitude is a measure of the star’s luminosity - the amount of energy radiated by the star every second. If you know the apparent magnitude and absolute magnitude, you can find the star’s distance. If you know apparent magnitude and distance, you can find the star’s luminosity. Some apparent magnitudes are given by the following table:

Sun	=	−26.8
Moon	=	−12.6
Venus	=	−4.4
Sirius	=	−1.4
Vega	=	0.00
faintest naked eye star	=	+6.5
brightest quasar	=	+12.8
limiting mag. of Hubble	=~	+28.0

Most apparently bright stars are intrinsically bright (luminous), but most nearby stars are intrinsically faint. If we assume we live in a typical patch of galaxy (Copernican principle), this means that most stars do not give off much light. The faintest stars have absolute magnitudes of 19 while the brightest stars have absolute magnitudes of −8. This translates to a huge range in luminosity!

The "Distance modulus" is equal to the apparent magnitude of a star minus its absolute magnitude. It can be calculated as follows:

app. mag. − abs. mag. = 5*log(distance in parsecs) − 5

Therefore, the ratio of two stars' luminosities is

L1/L2 = 10^{0.4*(abs.mag.1 − abs.mag.2)}

Remember that a more luminous star has an absolute magnitude that is less than a fainter star's absolute magnitude!

The following table compares the apparent magnitudes, absolute magnitudes, distances, and luminosities for various well-known objects:

Star	App.Mag.	Distance(pc)	Abs.Mag.	Luminosity(rel. to Sun)
Sirius	−1.4	2.7	1.4	23
Arcturus	0	11.0	−0.2	100
Vega	0	8.1	0.5	52
Spica	1	8.0	−3.4	1900
Barnard's Star	9.5	1.8	13.3	1/2500
Proxima Centauri	11.0	1.3	15.5	1/19000

1.45 Star Colors (Temperature)

STAR COLORS (TEMPERATURE)

The color of stars depends on their temperature - hotter stars are bluer and cooler stars are redder. We can measure a star's temperature by using different filters, which allow only a narrow range of wavelengths (colors) through. By sampling the star's spectrum at two different wavelength ranges ("bands"), we can determine if the spectrum is that for a hot, warm, cool, or cold star. Hot stars have temperatures of around 60,000 K while cold stars have temperatures around 3,000 K.

The B−V magnitude is a way of quantifying the star's color. The magnitude of the star through a standard blue filter is subtracted from the magnitude when taken through a "visible light" filter (which lets only greens and yellows through). A hot star has a B-V magnitude close to 0 or negative, while a cool star has a B-V magnitude close to 2.5.

View spectrum of a hot star
View spectrum of a cool star

For more information, see The Magnitude System, Spectral Classifications.

1.46 Stellar Composition

STELLAR COMPOSITION

We determine the composition of stars through spectroscopy - breaking the starlight up into individual colors and noting what absorption (or sometimes, emission) lines are present. We generally see absorption lines similar to those found in our Sun - hydrogen and helium with traces of other elements. From these absorption lines we learn some important things in addition to the star's composition:

1. Structure of stars: Stars are a hot dense body producing a continuous spectrum, covered by cooler thin gas.
2. The physics we use on Earth works everywhere else in the Universe! The hydrogen spectrum is same in the Sun, stars, distant galaxies and quasars. All absorption lines seen in celestial objects can be seen in laboratories on Earth. The charge and mass of the electron and proton are same everywhere we look. Physical laws are the same everywhere.
3. Since light has a finite speed, the light we receive from far away sources tell us how they appeared long ago. Since we see spectrae that can be explained with current terrestrial physical laws, it is easy to conclude that our physical laws have remained unchanged throughout time!

1.47 Velocities of Stars

VELOCITY OF STARS

We determine the velocity of stars by using the Doppler effect. A motion causes a wavelength shift, with the amount of shift dependent upon the speed of the object relative to the observer. A red shift means that the object is moving away, whereas a blue shift means that the object is moving toward you. This only gives us an object's speed along the line of sight. To get the tangential speed, we need to measure the angular speed of the star across the sky and determine the star's distance (both of which are much more difficult to do).

View picture

1.48 Stellar Masses

STELLAR MASSES

To determine the masses of stars we use spectroscopic binaries and Kepler's third law. Since stars have about the same mass (within a factor of 20), they both orbit around a center of mass that is significantly different from one of the star's centers. The center of mass is the point where:

(mass star 1) * (C.M. distance 1) = (mass star 2) * (C.M. distance 2)

or the point they would be balanced upon if the stars were on a stellar seesaw. A heavy star is closer to the center of mass than a light star. Kepler's 3rd law states that:

(mass 1 + mass 2) = a^3/P^2

if we use solar masses, A.U. for the distance from the center of mass 'a' and years for the orbital period 'P'.

A few rules assist in determining stellar masses:

1. Stars stay on opposite side of center of mass from each other.
2. The heavy star moves slower than the light star.
3. The center of mass is also the point where the mass times velocity of one star is equal to the mass times velocity of the other.

View picture

1.49 Stellar Radii

STELLAR RADII

To determine the radii of stars we can use eclipsing binaries. We can plot light curves - plots of brightness vs. time. An example light curve of a circular orbit seen edge-on with a small hot star and a large cool star is provided in the following image:

View picture

From analyzing such light curves, astronomers can make educated guesses as to the relative sizes of stars in a multiple system.

1.50 Temperature Dependence

TEMPERATURE DEPENDENCE

The strength and wavelength of absorption lines varies between stars. Some stars have strong (dark) hydrogen lines, while other stars may have no hydrogen lines but strong calcium and sodium lines. However, their relative abundances are not substantially different. Instead, temperature effects, in particular the temperature of the photospheres, cause the observed variations. The following effects of temperature can be observed:

1. To absorb a photon of a certain energy, an electron needs to be at the right energy level. If hydrogen atoms are heated to high temperatures, atomic collisions can cause their ionization, resulting in no absorption. If the star's temperature is too low, then there are not many electrons in the second energy level - most are in the ground state because there are not that many atomic collisions. To produce absorption lines in the visible spectrum, we need electrons in the 2nd energy level.
2. Hydrogen lines are strong for temperatures ranging from 4,000 to 12,000 K. Helium lines are strong for temperatures of 15,000-30,000 K. Calcium lines are strong for temperatures of 3000-6000 K.
3. The strength of lines is sensitive to temperature. Cross-referencing elements' line strengths provides accurate temperature measurements (within 20-50 K). Stars are not perfect thermal radiators so a continuum spectrum gives only a rough temperature (within a few hundred Kelvin).

1.51 Spectral Types

SPECTRAL TYPES

Originally, spectral types were based on hydrogen absorption line strength. A-type was the strongest, B-type the next strongest, F-type next, etc. Originally there was the whole alphabet of types, based on hydrogen line strengths, but then astronomers discovered that the lines depended on temperature. After some rearranging and merging of some classes we now have OBAFGKM classes ordered by temperature. Each class is subdivided into 10 intervals, e.g., G2 or F5, with 0 hotter than 1, 1 hotter than 2, etc.

For more information on the spectral classification of stars, see Spectral Classifications.

1.52 Hertzsprung-Russell (H-R) Diagram

HERTZSPRUNG-RUSSELL DIAGRAM

The Hertzsprung-Russell diagram (also called a color magnitude diagram) attempts to look for correlations between stellar properties. By plotting luminosity vs. temperature, Hertzsprung and Russell independently found a surprising correlation for 90% of the stars.

The stars lying along a diagonal strip (such as our Sun) are termed to be in the "main sequence". Luminous stars are easier to observe, but more rare than faint stars (which are difficult to see).

There also appears to be a correlation between a star's mass and its luminosity, with the empirical relation:

    Luminosity = constant * Mass ^ 3.5.

The H-R diagram below is plotted for all stars visible to the naked eye, plus all stars within 25 parsecs.

View example H-R diagram

1.53 Spectral Type to Distance

Astronomers can use the correlation between main sequence luminosity and temperature (spectral type) to get distances. The following steps are required:

1. Determine spectral type (from spectroscopy) and measure the star’s flux.
2. Use a calibrated main sequence star to get luminosity (the Hyades cluster in the constellation of Taurus is a commonly-used calibrator).
3. Use the Inverse Square Law for Brightness to get distance.

1.54 Cluster Main Sequence

Astronomers can use a color magnitude diagram for an individual cluster, and compare it with a similar diagram for a cluster to which the distance is known. By making adjustments for the cluster’s age and composition differences in stars (called "main sequence fitting"), the intrinsic luminosity for stars in the cluster can be determined. From this and their apparent magnitudes, their distances can be determined.

1.55 Stellar Structure

The following chapter is based on information provided by Nick Strobel. It has been edited, and included with "The Digital Universe" with the kind permission of the author.

STELLAR STRUCTURE

```
* The Sun’s Power Source
    + Timescale
    + Chain Process
    + Mass Sensitivity
    + Neutrinos
    + Solar Neutrino Problem
* Interior Structure of Stars - How do we know what they’re made of?
    + Mathematical Models
        1. Temperature
        2. Pressure
        3. Mass Density
    + Equation of State
    + Hydrostatic Equilibrium
```

1.56 The Sun's Power Source

The Sun's Power Source

The Sun produces a lot of light every second and it has been doing that for billions of years. How does it or any other star produce so much energy for so long?

The first basic question about the Sun is what powers it? It puts out A LOT of energy every second, equal to 8*10^19 five Megawatt power plants (nuclear or hydroelectric) on the Earth. Our best power plants now can produce around five Megawatts of power. Another way to look at this is that the Sun puts out every second the same amount of energy as 2.5*10^12 of those five Megawatt power plants would put out every year. That's over two trillion - numbers like that on the Earth only come up when talking about the national debt!

So, what powers our Sun and other stars? First, we rule out other likely candidates. Chemical reactions are not very likely. The most efficient chemical reaction is combining two hydrogens to one oxygen to make a water molecule plus some energy. Such a reaction has a very small efficiency (something like 1/10000000 of one percent). The efficiency means the net amount of energy returned from the reaction after the expenditure of energy getting the reaction to happen in the first place. To find out how long the Sun would last we would need to find out how much energy the Sun has stored in its account and know how fast it makes withdrawals on its account. The amount of time it would last would be the energy stored divided by the rate of withdrawal. The water reaction would only power the Sun for about 10,000 years.

We need a reaction with a higher efficiency. How about the ultimate in efficiency - a matter-antimatter reaction with 100% efficiency. Such a reaction could power the Sun for 10^13 years. Unfortunately, there are problems with this because the number of heavy particles in the Sun must stay the same and very soon the matter-antimatter reaction would violate that rule.

How about gravitational settling? This is a fancy way of referring to the converting of the potential energy of the falling layers to kinetic energy. When you hold a rock above the ground it has stored energy ("potential energy"--it has the potential to do some work). The stored energy is released as you let it fall. The rock gets kinetic energy because it is moving. Kinetic energy can heat things up. This is what would happen to the layers of the Sun if they were to fall inward toward the center of the Sun. The gas would be compressed and, therefore, would heat up. In addition to the expected heating the gas would also radiate. This was the idea physicists strongly argued for until the beginning of this century. This gravitational energy (with an efficiency of 1/10000 of one percent) could power the Sun for 3*10^7 years--a nice long time except for the nagging but ever louder criticism of the biologists who needed more time for evolution to occur and the geologists who preferred the idea of unlimited age for the Earth but would stomach something like a few billion years. A good article on the age-of-the-Earth debate is in "Scientific American" for August 1989, pages 90-96. Eventually physicists had to change their minds about the age of the Sun (and Earth) as radioactive dating (something physicists believe is correct) indicated a 4.6 billion year age for the solar system and, therefore, the Sun. It was the fact that the Sun could not last long enough being powered by gravitational contraction or settling that motivated the search for nuclear power sources.

We are left with nuclear reactions as the only thing left to power the Sun. There are two types possible: fusion and fission. Fission produces energy by breaking up massive particles like uranium into less massive particles like helium and lead. Fusion produces energy by fusing together light particles like hydrogen into more massive particles like helium. Atomic power plants and the atom bomb use fission to get the energy. Stars and hydrogen bombs use fusion. To get positively charged particles to fuse together, the electrical repulsion must be overcome, since like charges repel and opposite charges attract. Once the positively charged particles are close enough together (within several 10^{-13} centimeters of each other), the strong nuclear force takes over and is much more powerful than the electric force. The nuclei stick together. To get those particles close enough together requires high temperatures and high densities - something that occurs naturally in the cores of stars.

Fusion takes light particles whose combined mass is more than the resulting fused massive particle. The mass that is given up to form the massive particle is converted to energy via Einstein's equation $E=mc^2$. That tells you how much energy (E) can be made from matter with mass m. Remember that c is the speed of light and it's squared so a little bit of mass can make a lot of energy. An example for fusion is the fusion process in the cores of main sequence stars that takes four hydrogen nuclei (protons) each with the mass of one proton and fuses them to form a helium nucleus (two protons and two neutrons) that has the mass of 3.97 times the mass of one proton. An amount of mass equal to 0.03 times the mass of one proton was given up and converted to energy equal to $0.03*(\text{mass one proton})*c^2$. The efficiency of this reaction is about 4/5 of one percent. The Sun could last for about 10 billion years on hydrogen fusion in its core. This is plenty long enough to satisfy the modern geologists.

View proton-proton fusion reaction

The next basic question is: why does nature use the long complicated proton-proton chain (or the longer carbon-nitrogen-oxygen chain) for fusion? Nature uses a three step chain process to fuse four protons to make one helium nucleus for the proton-proton chain process. Wouldn't it be much simpler if four protons would collide simultaneously to make one helium nucleus? Simpler, but not very likely is the answer. Getting four objects to collide simultaneously is very hard to do - the chances of this happening are very very small. The chances of this type of collision are too small to power the Sun, so nature has found a trickier scheme. The chances of two particles colliding and fusing is much higher, so nature slowly builds up the helium nucleus.

Nuclear fusion is something of a holy grail for utility companies because it produces no nasty waste products and has the potential of getting more energy out of it than you put in. Unfortunately, the conditions to get fusion to happen are very extreme by our standards. We've been only able to tap the fusion process with the hydrogen bomb, but that's a one shot deal. The hydrogen bomb still needs an atomic bomb trigger to get the extreme temperatures needed for the fusion process. At least we can get the waste product of the Sun's fusion process for free with solar power collectors. The Sun can have a controlled fusion process and not blow up all at once because of the hydrostatic equilibrium "thermostat".

Hydrostatic equilibrium is the balance between the thermal pressures from the heat source pushing outwards and gravity trying to make the star collapse to the very center. The nuclear fusion rate is very sensitive to temperature. It increases as roughly T^4 for the proton-proton chain and even more sharply (T^{15}) for the carbon-nitrogen-oxygen chain. So a slight increase in the temperature causes the fusion rate to increase by a large amount and a slight decrease in the temperature causes a large decrease in the fusion rate.

Now suppose the nuclear fusion rate speeds up for some reason. Then the following sequence of events would happen:

1) Thermal pressure would increase, causing the star to expand
2) The star would expand to a new point where gravity would balance the thermal pressure, but
3) Expansion lowers temperature in the core - nuclear fusion rate slows down
4) Thermal pressure drops and the star shrinks
5) Temperature rises and nuclear fusion rate increases
6) Stability is reached between nuclear reaction rates and gravity.

A similar type of scheme would occur if the nuclear fusion rate were to slow down for some reason. The fusion rate stays approximately constant for stars that are fusing hydrogen to make helium plus energy in the core. Once the hydrogen fuel in the core has been used up, hydrostatic equilibrium can no longer stabilize the star. What happens next will have to wait until we talk about stellar evolution.

Here’s a summary of fusion in outline form along with some discussion of a particle produced from fusion - the neutrino.

A. TIMESCALE

We need an energy source that lasts a long time. Nuclear fusion causes lighter nuclei to fuse together to form a heavier nucleus + energy. The sum of the light nuclei masses = heavy nucleus mass + energy/c^2 since E=mc^2. The "c" is the symbol for the speed of light.

B. CHAIN PROCESS

A chain process is used to fuse hydrogen to helium. Chain processes are much more probable than "all at once" processes. The proton-proton and carbon-nitrogen-oxygen reactions are chain processes.

View proton-proton fusion reaction

C. MASS SENSITIVITY

The reaction rate is sensitive to mass, with more mass resulting in more reactions. Main sequence type stars are hydrogen fusion powered.

D. NEUTRINOS

A neutrino is a massless (or nearly massless) particle travelling at or near the speed of light. It is produced in nuclear reactions. It has a low probability of interaction with normal matter and can pass directly out of the core of a star. If we could catch neutrinos, we could find out what it was like in the Sun’s interior just 8.3 minutes ago. Photons, on the other hand, take about a million years to percolate out to the surface. We can increase the odds of neutrino detection by using a large amount of material that reacts with neutrinos in a certain way (for example, chlorine changes to radioactive argon, and gallium and water molecules give off flashes of light when interacting with neutrinos).

E. SOLAR NEUTRINO PROBLEM The problem is that in experiments, we get only 1/3 the expected number of neutrinos! Possible reasons:

1. Fusion may not be the Sun's power source (not supported by other observations).

2. The experiments may not be calibrated correctly. However, it is unlikely that all the carefully tuned apparati are tuned in the same wrong way.

3. The nuclear reaction rate could be lower than what our calculations say. This is possible, but many people have checked and re-checked the physics of the reaction rates.

4. Neutrinos change into other types of neutrinos that don't interact with the detection material during their flight from the Sun to Earth (this idea is gaining more and more advocates). This theory requires the neutrino to have some mass. If the neutrino has mass, then it cannot travel at the speed of light, but it can get close. Also, a neutrino with mass has important consequences for the evolution of the universe (more about that later).

1.57 Mathematical Models

One of the ways in which we can determine the state of the interior of a star is through the use of mathematical models. These are sets of equations which describe how things work. Since the interior is gaseous, the equations are relatively simple. The equations are based on three main parameters:

1. Temperature - the average kinetic energy of gas particles.
2. Pressure - force per unit area. Hot gas wants to expand, which causes pressure.
3. Density - mass per unit volume. Gas can be compressed to smaller volumes, resulting in larger densities

1.58 Equation of State

We can deduce the structure of a star's interior by using the Equation of State. This is an equation relating density, pressure, and temperature. The equation of state for a gas is relatively simple:

pressure = constant * density * temperature / molecular weight

For hydrogen, the molecular weight is very close to 1, and for helium the molecular weight is very close to 4. For a gas made of different types of atoms, the molecular weight is the weighted mean of the different atomic types, taking into account their relative proportions.

1.59 Hydrostatic Equilibrium

Conditions within a star’s interior can be further hypothesized through studies of the conditions of hydrostatic equilibrium. A star is held together by gravity, as everything is compressed towards the center. Thermal pressure, on the other hand, tries to expand the star’s layers to infinity. In any given layer of a star there is a balance between thermal and gravitational pressure. This is known as hydrostatic equilibrium.

Deeper layers of a star have more gravitational compression and hence require more thermal pressure to reach an equilibrium. From these studies, we find that the temperature of the core of our Sun must lie between 8 to 28 million degrees Celsius and have densities of between 10 to 130 grams per cubic centimetre. As stars age, these numbers increase.

1.60 Stellar Evolution

Much of the following chapter is based on information provided by Nick Strobel. It has been edited, and included with "The Digital Universe" with the kind permission of the author.

STAGES IN THE LIFE OF A STAR

* The History of a Star
    + Timescales
    + Stages
        1. Giant Molecular Cloud
        2. Protostar
        3. T-Tauri
        4. Main Sequence
        5. Subgiant, Red Giant, Supergiant
        6. Core fusion
        7. Red Giant, Supergiant
        8. Remnant
* Stellar Remnants
    + Brown Dwarfs
    + White Dwarfs
    + Neutron Stars
    + Black Holes

1.61 The History of a Star

The History of a Star

The stages and timescales involved in the development of a star are mostly dependent upon the star’s original mass (with a bit of dependence on the star’s composition as well). In general, massive stars evolve more quickly than lighter stars.

STAGES IN THE DEVELOPMENT OF A STAR

1. Giant Molecular Cloud

This is a large dense cloud of cold dust and gas. It typically contains 100,000 to a few million times the mass of our Sun. As a result of gravity and/or shock waves, portions of the cloud begin to collapse into pockets 10 to 100 times more massive than our Sun.

2. Protostar

This is a large gas "clump" collapsing and heating up in the center as it does so. Gravitational energy is being converted into heat. It gets hot enough to glow red (2000 - 3000 K), but a "cocoon" of gas and dust prevents the visible light from escaping. Infrared and microwave radiation, however, can pass through the cocoon and is detectable by Earth-based instruments.

3. "T-Tauri" starlike object

Eventually the star becomes hot enough at its center for nuclear fusion to initiate. When this is done, strong stellar winds eject a lot of material from the star and the cocoon of gas and dust is blown away. The young star can then be detected at visible wavelengths.

4. Main sequence star

A typical star spends about 90% of its lifetime as a main sequence star. It is in the process of fusing hydrogen to helium in its core. The star remains in a stable state because of the processes of hydrostatic equilibrium.

5. Subgiant, Red Giant, or Supergiant

As the hydrogen fuel in the core of a star is converted into helium, a point is eventually reached where insufficient hydrogen remains to undergo fusion. The hydrostatic equilibrium of the star is upset, the core shrinks, and the star begins to collapse. As it does so, the outer layers of the star heat up and fusion begins there. Fusion in the outer layers is very rapid, so the energy output of the star increases and the gas envelope surrounding the core puffs out. Since the surface area of the star is now much larger, the energy is spread out and each square centimetre of the star’s surface becomes a cooler, red color. These Red Giant stars can eject a lot of mass through stellar "winds". This mass can take the form of a planetary nebula. Note that although a red giant is large in terms of linear size, it is less massive than the main sequence star it came from.

6. Core fusion

As the core continues to shrink, the temperature rises until elements heavier than hydrogen can begin to undergo fusion. Typically, helium begins to fuse. In low mass stars, the onset of helium fusion can be very rapid, producing a burst of energy known as a "helium flash". Eventually, the fusion process stabilizes and the star begins a second "life" fusing helium. More energy is released now than when the star was on the main sequence, so the star is bigger. However, the length of time that the star can remain stable at this stage is much less than when it was burning hydrogen.

7. Red Giant, Supergiant

Eventually, the supply of helium in the core runs out as did the hydrogen. If the star is massive enough, stages 5 and 6 will repeat as even heavier elements begin to fuse. The process cannot go on indefinitely, however. When iron nuclei are produced as a byproduct of the fusion process, the star has reached a "dead end". More energy is required to get iron nuclei to fuse together than is released in the fusion process, so the star virtually runs out of nuclear fuel.

8. Remnants

Now that nuclear reactions can no longer produce the heat required to offset the crush of gravity, the star begins its final collapse. What happens next is highly dependent on the mass of the star. See the section on Stellar Remnants for further information.

1.62 Stellar Remnants

Stellar Remnants

Once the nuclear fuel in a star has been exhausted, it can no longer produce the heat required to offset the crush of gravity. What happens next is highly dependent on the original mass of the star (its mass while it was on the main sequence). The following chart helps to summarize the possible stellar remnants (the original mass is given in solar masses, where 1.0=the mass of our Sun):

Original Mass	Remnant
0.01 to 0.08	Star collapses relatively gently. A brown dwarf star is created.
0.08 to 0.25	Star collapses relatively gently. A white dwarf star with a core rich in helium is created.
0.25 to 8.0	Star collapses relatively gently. A white dwarf star with a core rich in carbon and oxygen is created. The final mass of the white dwarf is less than 1.4 solar masses.
8.0 to 12	Star collapses relatively gently. A white dwarf star with a core rich in oxygen, neon, and magnesium is created. The final mass of the white dwarf is less than 1.4 solar masses.
12 to 40	Star collapses with a tremendous supernova explosion. A neutron star is created, with a mass between 1.4 and 3.2 times that of our Sun.
>40	Star collapses with a tremendous supernova explosion. A black hole is created, with a mass greater than 3.2 times that of our Sun.

1.63 The Interstellar Medium

Much of the following chapter is based on information provided by Nick Strobel. It has been edited, and included with "The Digital Universe" with the kind permission of the author.

Briefly, the interstellar medium is the "stuff between the stars". It is believed to constitute 10-15% of the visible mass of the Galaxy, and is composed of 99% gas and 1% dust. Since stars are formed from the interstellar

```
medium, and it affects the starlight passing through it, the ISM is of great
interest to astronomers.

                              INTERSTELLAR MEDIUM

          + Dust
              1. Extinction
              2. Reddening
          + Gas
              1. H-II Regions
              2. H-I Regions
              3. Molecular clouds
```

1.64 Dust in the Interstellar Medium

```
DUST

Dust is considered to consist of particles smaller than 0.005 millimetres in
diameter.  Dust in the interstellar medium is usually composed of water ice,
graphite (carbon), or silicon in highly flattened flakes or needles.  It has
several effects on light passing through it:

1.  Extinction

Extinction is the dimming of starlight at all wavelengths.  In 1930, R. J.
Trumpler plotted the angular diameter of clusters vs. the distance to the
cluster derived from the Inverse Square Law. If the clusters are believed to
have nearly the same linear diameter, the the angular diameter should be
equivalent to their linear diameter divided by their distance.  Instead, he
found a systematic increase of the linear size of the clusters with distance.
He couldn't believe that nature had put the Sun at a special place where the
size of the clusters was the smallest.  Instead, he felt that the Sun is in a
typical spot, but that more distant clusters appear dimmer than would be
predicted by a straightforward application of the Inverse Square Law.  This
would imply that dust in the interstellar medium makes the stars appear fainter
(further away) than they really are.

2.  Reddening

Though the interstellar medium causes the intensity of starlight to be reduced
at all wavelengths, in fact the extinction depends somewhat upon the
wavelength, with blue light being scattered more than red light.  Since less
blue light reaches us, more distant objects appear redder than they should.
These theoretical deductions were supported by scientific measurements obtained
by Trumpler which showed that a given spectral type becomes increasingly redder
with distance.  The same process makes the Sun appear redder when it is near to
the horizon, though in this case the reddening effect is caused by dust in our
atmosphere instead of that between the stars.
```

1.65 Gas in the Interstellar Medium

The gas in the interstellar medium is composed of about 90% hydrogen and 10% helium, with traces of other elements. We are able to observe the hydrogen in ionized, neutral atomic, and molecular forms.

Gas in the interstellar medium is observed when looking at the spectral lines of a binary star. We see narrow absorption lines that do not move and broader lines that shift as the stars orbit each other. The broad lines result from absorption in the star's atmosphere whereas the narrow lines indicate absorption from the ISM.

Astronomers have detected several regions in space with differing characteristics of interstellar gas:

1. Ionized hydrogen atoms (H-II regions)

Ultraviolet light from hot O & B (spectral type) stars is absorbed by hydrogen gas and re-emitted mostly at visible wavelengths, primary 6563 Angstroms. Each ultraviolet photon produces a visible photon. Spectrae from these regions are much simpler than stellar spectrae.

2. Neutral hydrogen atoms (H-I regions)

Hydrogen in space is in the ground state. The electron in the atom can either be spinning in the same or the opposite direction as the proton. If it is spinning in the opposite direction, the energy state is slightly lower. Normally, the spin directions are opposite (as it is the lower energy state), however sometimes atoms collide and the spins are re-aligned. Eventually (every few million years, on average), the electron flips its spin to get back into the lowest energy state. When it does so, a photon with a wavelength of 21.1 cm is emitted. This is a rare occurrence, but with the large amounts of hydrogen in space, a noticeable amount of 21 cm radiation has been detected. The Milky Way has about 3 billion solar masses of H-I gas, most in the outer regions of the galaxy. Since the long wavelengths are not impeded by interstellar dust, H-I regions can be easily detected.

3. Molecules

Several molecular clouds have also been detected. Most of these clouds consist of hydrogen gas (H2) or carbon monoxide (CO), but over 100 other molecules have been detected as well. The molecular clouds can range in size from just a few solar masses to over a million, with radii ranging from a few to over 100 parsecs. Our Milky Way has about 2.5 billion solar masses of molecular gas, with about 70% of it in a ring 4000-8000 parsecs from the centre. There is not much molecular gas between 1000-3000 parsecs from the centre. Most of the hydrogen gas is clumped in the spiral arms within the disk and stays within 120 parsecs of the disk midplane.

1.66 Galaxies

Much of the following chapter is based on information provided by Nick Strobel. It has been edited, and included with "The Digital Universe" with the kind permission of the author.

GALAXIES

```
* Galactic Structure of the Milky Way
     + Period-Luminosity Relation for Variable Stars
         1. Cepheids
         2. RR-Lyrae
     + Our Location
     + Our Motion
     + Deriving the Galactic Mass from the Rotation Curve
     + Spiral Arms
         1. Density wave theory
         2. Transient spirals
         3. Supernova creating spiral arms
     + Populations of Stars
         1. Population II
         2. Population I
     + Galactic Center
* Other Galaxies
     + Types of Galaxies
         1. Ellipticals
         2. Spirals
         3. Irregulars
     + Positions
     + Distances to Galaxies
     + Masses of Galaxies
     + Origin of Galaxies
     + Clusters of Galaxies
         1. Local Group
         2. Virgo Cluster
         3. Coma Cluster
     + Superclusters
* Steps to the Hubble Constant: Distance scale ladder
     + Rung 1:  Distance to Sun
     + Rung 2:  Trigometric Parallax
     + Rung 3:  Spectroscopic Parallax
     + Rung 4:  Variable Star Period-Luminosity Relation
     + Rung 5a: Galaxy Luminosity vs. Another Bright Feature
     + Rung 5b: Luminosity or Size of Bright Feature
     + Rung 6:  Galaxy Luminosity and Inverse Square Law
     + Rung 7:  Hubble Law
```

1.67 Galactic Structure of the Milky Way

Galactic Structure of the Milky Way

View diagram of our Milky Way

Astronomers know that the Milky Way is disk-shaped with spiral arms. There is an elliptical bulge in the centre and a spherical halo that is more dense closer to the galactic center. But how do we know these things? The following observations were all required before astronomers could remark on the structure of our galaxy with any certainty:

A. PERIOD-LUMINOSITY RELATION FOR VARIABLE STARS

Some stars in the last stages of life pulsate by changing size. They try to have hydrostatic equilibrium, but thermal pressure lags behind the gravitational collapse. The expanding star overshoots its equilibrium point. Then, gravity catches up and contracts the star beyond its equilibrium point. The cycle continues. There are two main classes of variable stars of interest in determining the structure of our Milky Way - Cepheids and RR Lyrae variables. They are both extremely useful as "cosmic yardsticks" enabling astronomers to determine their distances from us.

B. OUR LOCATION

In 1918, Harlow Shapley used the period-luminosity relation for variable stars to find the distance to 93 globular clusters in very elliptical orbits around the center of our galaxy. He found that there was a strong concentration of globulars in the direction of the constellation Sagittarius. He took this concentration to indicate the center of our Milky Way and determined that we are located at a distance of about 10,000 parsecs from the centre. More modern estimates place us at a distance of about 8,000 parsecs.

C. OUR MOTION

Astronomers observed that globular clusters on one side of the celestial sphere have red shifts while globulars on the other side have blue shifts. In addition, stars and H-II regions near the Sun have small doppler shifts. They concluded that the disk of our galaxy is rotating in an organized fashion, while globular clusters are not. Today, we are able to use the 21 cm emission from neutral hydrogen (H-I regions) to map out the motion of this disk.

D. DERIVING THE GALACTIC MASS FROM THE ROTATION CURVE

By studying the amount of time it takes for stars and clusters to revolve around the centre of the Milky Way, scientists are able to calculate the approximate mass of our galaxy. However, when they first did so, they discovered a surprising fact - that the galaxy contains ten times more matter than can be seen. This came to be known as "dark matter", and it could be in the form of planets, brown dwarfs (too small to be a star), black holes, neutrinos with mass, or other exotic particles that we haven’t discovered in our laboratories yet. A lot of research is currently being done to discover the nature of this dark matter.

E. SPIRAL ARMS

Gas is compressed in spiral arms to form stars. O & B stars (very hot, massive, young stars) and H-II regions enhance the spiral outline. Lots of dimmer stars lie between the arms. Differential rotation of the arms means that they should wind up and last only about 500,000,000 years. But other galaxies have spirals that are not wound up, so that means that spirals are long-lasting and can sometimes occupy an entire galaxy. Astronomers are still not sure as to why this happens, but they have a few theories:

1. Density wave theory: This theory says that a spiral region of greater gravity concentrates stars and gas. The spiral regions rotate more slowly than the stars do (by about a half). Stars behind the region of greater gravity are pulled forward into the region and speed up. Stars leaving the region of greater gravity are pulled backward and slow down. Gas entering the spiral wave is compressed. On the downstream side of wave, there should be lots of H-II regions (star formation regions). Scientists have not yet proven that star formation regions exist in these

locations, nor have they been able to answer what forms the wave in the first place, or what maintains the wave.

2. Transient spirals: This theory states that spiral arms come and go. Computer simulations of galactic disks sometimes show spiral patterns appearing and disappearing.
3. Supernova creating spiral arms: Some astronomers believe that massive stars in a spiral pattern die and undergo a supernova explosion. This explosion compresses the surrounding interstellar medium, causing the formation of more stars, some of which undergo more supernova explosions, and so on. The shock waves keep the spiral arms visible. However, it is unclear as to why these chains of supernova explosions might occur in a spiral pattern.

F. POPULATIONS OF STARS

In 1944, Walter Baade discovered that stars generally fell into two basic groups:

Population I : These stars are young and rich in metals. They are mostly found in the disk component of our galaxy, and have a wide range of ages (0-10 billion years), with the youngest ones in the spiral arms.

Population II: These stars are older and poor in metals. They are mostly found in the spheroidal component (halo, bulge) of our galaxy and are generally more than 10 billion years old. Population II stars have large random velocities with highly elliptical orbits.

G. GALACTIC CENTER

The center of our galaxy was observed to be a strong radio source in the constellation Sagittarius (the source is known as Sagittarius A). The signal is believed to originate from rapidly moving charged particles in a strong magnetic field. X-rays are also seen in the immediate vicinity. Stars near the center are orbiting quite quickly, indicating that a mass more than one million times that of the Sun lies nearby. Astronomers expect that this may be evidence of a massive black hole formed by the mergers of stars and stellar remnants. It may also be simply a dense cluster of stars, such as has been identified in some other galaxies. We also see an expanding ring about 3000 parsecs from the center that astronomers are unable to provide a definite explanation for.

1.68 Other Galaxies

A galaxy is an organized system of tens of millions to trillions of stars, sometimes mixed with gas and dust and all held together by gravity.

A. TYPES OF GALAXIES

There are many different types of galaxies. Edwin Hubble classified several galaxies into one of a few different categories. He originally thought that the categories represented an evolutionary sequence, but we now believe that they do not.

1. Ellipticals - These galaxies have round or elliptical shapes. Star motion within the galaxies is quite random, so their elliptical appearance is not due to rotational flattening. There

is very little dust and gas between the stars, and as a result there is no new star formation or hot, bright, massive stars. Most elliptical galaxies are small and faint, and dwarf ellipticals may be the most common type of galaxy known.

2. Spirals - These galaxies exhibit a flattened disk with a spiral structure. The spiral arms can go all the way to the central bulge, or be attached to a bar that bisects the bulge (galaxies exhibiting characteristics of this latter case are often grouped into the subcategory "barred spirals"). There is more rotational motion than random motion, and there is more gas and dust between the stars than in elliptical galaxies. As a result, new star formation is occurring in the spiral arms.
3. Irregulars - Irregular galaxies show no definite structure. Some irregulars have lots of dust and gas, making star formation possible. Most galaxies of this type are quite faint.

B. POSITIONS

When astronomers first began to detect galaxies, they noticed that they seemed to avoid a 5 to 10 degree band along the mid plane of the Milky Way. This region was termed the "zone of avoidance". We now know that the missing galaxies are simply obscured by vast lanes of dust within our own galaxy.

C. DISTANCES TO GALAXIES

If we want to determine the luminosity and mass distribution of galaxies, we need to be able to determine their distances. The section entitled Steps to the Hubble Constant contains more information on how this can be done.

D. MASSES OF GALAXIES

Most galaxies contain large amounts of "dark matter". This is material which does not give off light, but has a noticeable gravitational effect. Some astronomers estimate that there may be ten times as much dark matter as there is visible matter.

E. ORIGIN OF GALAXIES

For a long time, astronomers have been trying to determine the origin of galaxies and their life cycle. The classical model on the origin of galaxies considers several possibilities:

1. A slowly rotating cloud of gas collapses and forms stars before the stars themselves flatten into a disk-shaped galaxy. As a result, an elliptical galaxy is formed.
2. A quickly rotating cloud forms a disk before most stars are created, so a spiral galaxy is formed. If the spiral galaxy is lacking a massive dark matter corona, a bar may form across the middle of it.

Several interesting galaxy structures may result from the collision of one or more galaxies. There is a great deal of computer simulation being done to try and explain how these structures might originate.

F. CLUSTERS OF GALAXIES

Stars cluster into galaxies, and galaxies cluster into even larger groups. Some of the more important such groups include:

1. Local Group - This is the group in which our Milky Way is a

member. It consists of 3 spirals, 2 ellipticals, 4 irregulars, and 8 dwarf ellipticals and is thought to be about 1,000,000 parsecs in diameter.

2. Virgo Cluster - This is an irregular cluster of hundreds of spiral, elliptical, and irregular galaxies, located between 15,000,000 and 18,000,000 parsecs from us. It is the closest large cluster, and we are actually "falling" towards it because of its large mass.
3. Coma Cluster - This is a huge cluster of thousands of galaxies (mostly elliptical and borderline spirals) in a large, regular shaped cluster 100,000,000 parsecs away. The ellipticals are in the central regions while the spirals lie on the outskirts.

G. SUPERCLUSTERS

Even galaxy clusters are not the largest groupings in the universe. There are also clusters of galaxy clusters, known as superclusters. They consist of tens to hundreds of clusters bound together into long filaments 100,000,000-300,000,000 parsecs long, 50,000,000-100,000,000 parsecs wide, and 5,000,000-10,000,000 parsecs thick. Between these filamentary superclusters are huge voids of very few galaxies. Astronomers do not know why these voids exit.

1.69 Steps to the Hubble Constant

Steps to the Hubble Constant

Why do we care so much about finding distances in astronomy? If we know the distance to a star, we can determine its luminosity and mass. We then can discover a correlation between luminosity, mass, and temperature for main sequence stars that our physical theories must account for. If we can measure the angular size of a star, we can then find its geometric size (how many kilometers in diameter it is). That gives us another clue to what is happening with the stars. Finding distances to stellar explosions like planetary nebulae and supernovae enables us to find the power needed to make the gaseous shells visible and how much was needed to eject them at the measured speeds. Stellar distances and distances to other gaseous nebulae are necessary for determining the mass distribution of our galaxy. We have been able to discover that most of the mass in our Galaxy is not producing light of any kind and is in a dark halo around the visible parts of the Galaxy (dark matter).

Finding distances to other galaxies enables us to find their mass, luminosity, and star formation history among other things. We’re better able to hone in on what is going on in some very active galactic cores and also how much dark matter is distributed among and between galaxy cluster members. From galaxy distances, we’re also able to answer some cosmological questions like the large-scale geometry of space, age of the universe, and whether or not the universe will continue expanding. This is only a quick overview of the reasons for distance measurements and is by no means an exhaustive list of reasons why we care about distance measurements.

Now let’s take a look at the distance scale ladder. The bottom foundational rung of the ladder is the most accurate and the most certain of all the distance determination methods. As we climb upward, each rung depends on the previous rung and distances become less certain.

RUNG 1

The Earth and distance to the Sun - We use radar reflections from Venus and the angular separation from Sun to get the Astronomical Unit (AU). We are able to use such methods to find distances out to 50 AU.

RUNG 2

Geometric Methods - We can determine trigonometric parallax to nearby stars by using their angular shift throughout the year and the Astronomical Unit. We can find the distances to nearby clusters (like the Hyades or Pleiades) via trigonometric parallax or the "moving clusters" method (another geometric method). Once the cluster's distance has been measured, we calibrate the cluster's main sequence in terms of absolute magnitude (luminosity). Geometric methods allow us to find distances out to 100 parsecs.

RUNG 3

Main Sequence fitting and Spectroscopic Parallax - We can find the spectral type of star and measure its flux. We then use a calibrated color-magnitude diagram to get its luminosity and from that calculate its distance. If we plot the cluster's main sequence on a color-magnitude diagram using the apparent magnitude instead of the absolute magnitude, we can determine how far the unknown main sequence needs to be shifted to match the calibrated main sequence. The age of a cluster affects its main sequence, as an older cluster only has fainter (slower- burning) stars left on the main sequence. As well, stars on the main sequence brighten at a constant temperature as they age, so they move vertically upwards on a Hertzsprung-Russell Diagram. We can model main sequence evolution to get back to a zero-age main sequence. If we assume that all zero-age main sequence stars of a given temperature (mass) start at the same luminosity, this technique allows us to find distances out to 50,000 parsecs.

RUNG 4

Variable Star Period-Luminosity Relation - We can find Cepheids and/or RR-Lyrae variables in stars clusters with their distances known through main sequence fitting. Or, we can use the more direct Baade-Wesselink method, which uses the observed expansion speed along the line of sight from doppler shifts with the observed angular expansion rate perpendicular to the line of sight. Such observations allow us to find distances out to 4,000,000 parsecs with ground-based telescopes, or 40,000,000 parsecs with the Hubble Space Telescope.

RUNG 5A

Galaxy Luminosity vs. Another Bright Feature - First, we find Cepheid variables in other nearby galaxies to get an estimate of the distance. We then use the galactic flux and the inverse square law of brightness to get the galactic luminosity. We either find the geometric size of H-II regions in spiral and irregular galaxies and calibrate this size with the galactic luminosity, or calibrate the correlation between the width of the 21cm line emitted by neutral hydrogen and the spiral galactic luminosity. The latter correlation is known as the Fisher-Tully relation. Elliptical galaxies have a correlation between their luminosity and their velocity dispersion in a relation known as the Faber-Jackson law. Distances of up to 100,000,000 parsecs can be determined with this method.

RUNG 5B

Luminosity or size of Bright Feature - First, we find Cepheid variables in other nearby galaxies to get an estimate of the distance. We can calibrate a number of different things from this:

1. Supernova maximum luminosity in any type of galaxy.
2. Globular cluster luminosity function in elliptical galaxies.
3. Blue or Red Supergiant stars relationship in spiral and irregular galaxies.
4. Maximum luminosity and rate of decline relationship of novae in ellipticals and bulges of spirals.
5. Planetary nebulae luminosity function in any type of galaxy.

Distances of up to 150,000,000 parsecs can be determined with this method.

RUNG 6

Galaxy Luminosity and Inverse Square Law - We can calibrate the Hubble law using the methods outlined in Rung 4 for nearby galaxies and Rung 5 for more distant ones. If those Rung 5 galaxies are like nearby ones (or have changed their luminosities in a known way), then using the flux and/or angular size and estimated luminosity and/or geometric size, we can find their distance. We need to consider the effect on the measured velocities caused by the Milky Way falling into the Virgo Cluster. We also find the galaxy cluster luminosity function. Hubble's law relates a galaxy's recession (expansion) speed with its distance by the relation:

speed=H0*distance

where H0 is Hubble's constant. Astronomers are perpetually trying to determine the value of H0 more accurately (see Our Expanding Universe for further details). This method allows us to determine distances out to 500,000,000 parsecs.

RUNG 7

In theory, if Hubble's constant has been determined and his law verified, it will be useful to determine distances to all far-off galaxies. We could use it to determine the geometry of the universe.
At the moment, Rung 4 is the critical step in the distance scale ladder. With the Hubble Space Telescope, we can use the Cepheid period-luminosity relationship out to distances ten times further than to what we can achieve on the ground. Some interesting results have been obtained. See Our Expanding Universe for more information.

1.70 Cosmology

Much of the following chapter is based on information provided by Nick Strobel. It has been edited, and included with "The Digital Universe" with the kind permission of the author.

COSMOLOGY

```
    * Observations and Some Implications
         + Universe has mass
         + Dark skies - Olbers' paradox
         + Uniform Universe
             1. Cosmological Principle
             2. Perfect Cosmological Principle
         + Universe without a center
         + Cosmic Microwave Background Radiation
         + Cosmic abundance of Helium and Hydrogen
    * Fate of Universe
         + Mass dependence
         + Critical density
         + Changing Hubble constant
         + Cosmic inventory
```

1.71 Observations and Some Implications

```
Observations of the Universe and Some Implications

A. THE UNIVERSE HAS MASS

Newton knew that if the universe is eternal and static (no net pattern of
motion), then the fact that there is matter in the universe must mean that
gravity will eventually cause the universe to collapse in upon itself.  If the
universe had existed for an infinite amount of time before the present, this
should already have happened (but it clearly has not).  There are at least
three ways to resolve this paradox, one or more of which may be true:
   1. The universe is infinite in volume and mass.
   2. The universe is expanding fast enough to prevent collapse.
   3. The universe is not eternal (it has a beginning and an end).
 Newton chose to think that the universe was infinite in volume and mass.

B. DARK SKIES - OLBERS' PARADOX

The sky at night is dark.  If the universe is infinite, eternal, and static,
then eventually any line of sight should end upon the surface of a star.
Therefore, the sky should be as bright as the surface of the Sun all the time!
This is known as Olbers' paradox after Heinrich Olbers who popularized the idea
in 1826 (though Kepler is known to have advanced the same idea in 1610).

There have been many possible explanations to try and resolve the paradox.
Some of the more useful ones include:

     1. There is too much dust to see distant stars.
     2. The Universe does not have an infinite number of stars.
     3. The distribution of stars is not uniform.  There could be an
        infinite number of stars, but they hide behind each other so that
        only a finite angular area is subtended by them.
     4. The Universe is expanding, so distant stars are red-shifted
        into obscurity.
     5. The Universe is young, and light from the distant stars hasn't
        reached us yet.

Though the first solution seems good at first glance, it cannot be the right
one.  Even if dust were in the way, its blocking of the light would heat it up
```

so that it would radiate energy.

The second solution has flaws as well. Even if the universe did not contain an infinite number of stars, there would still be enough to light up the sky.

The third solution may be partially responsible. We just do not have a really good idea of the distribution of stars in the universe. But even if the third solution is possibly correct, the last two solutions are definitely correct and contribute to a solution of the paradox. Since the Universe is believed to have started in the "Big Bang" about 15 billion years ago, any light more than 15 billion light years distant cannot have reached us.

C. UNIFORM UNIVERSE

Many astronomers believe that the universe is uniform on large scales (billions of light years). This is more of an assumption than an observation. It makes mathematical models of the universe easier to deal with. There are two variations of this theory:

1. Cosmological Principal - This states that the universe is homogeneous (there is no preferred observing position) and isotropic (we can see no difference in the general structure of the universe as we look in different directions). This principal is a Copernican idea, stating that we are not situated in a particularly special place in the Universe. The Cosmological Principal also assumes that the universe can change (evolve) with time.
2. Perfect Cosmological Principal - This theory encompasses the Cosmological Principal with the modification that the universe does not change (evolve) with time. If the universe is expanding, then new matter must be being created. Many astronomers do not believe in such a concept, since it would violate the conservation of mass law (even though the amount of "spontaneous" matter creation would be quite small - one atom of hydrogen per cubic metre every billion years).

D. UNIVERSE WITHOUT A CENTER

Since the universe is expanding in three dimensions and we are located inside this expansion, the center of expansion of the universe must remain undefined. Since every particle in the expanding universe sees every other particle moving away from it, each particle appears to itself to lie at the center of the universe.

E. COSMIC MICROWAVE BACKGROUND RADIATION

Cosmic background radiation was predicted by George Gamov in 1948 and first observed by Arno Penzias and Robert Wilson in 1965. With the COBE satellite in 1989, astronomers detected a microwave emission coming from everywhere in the sky that has the same intensity and spectral character as a thermal source at 2.735 +/- 0.06 K. This observation tells us a great deal about the early universe.

The further into space we observe, the further back in time we look. After the Big Bang, the density of radiation in the universe was so great that it dominated the expansion and conditions of the universe for 10,000 years. The early universe was hot and opaque, consisting of freely moving electrons, protons, and neutrons that scattered photons about. This made the dense gas glow with a continuous spectrum. As the universe expanded, it cooled off and eventually reached a point where electrons and protons could combine to form hydrogen atoms. When this happened, photons could travel much further distances without running into other particles. In effect, the universe (now at a temperature of about 3000 K) became transparent. It is from this time (700,000 years after the Big Bang) that the highly redshifted photons are reaching our radio telescopes as background radiation.

F. COSMIC ABUNDANCE OF HELIUM AND HYDROGEN

Hydrogen, being the simplest element, was formed first in the early universe. Since conditions were hot and dense, fusion was possible. From 1 to 3 minutes after the Big Bang, the universe was like the core of a star with 4 hydrogen atoms creating helium and releasing energy in the process. A small amount of lithium-7 was also produced during this period.

By the time the universe had expanded so that fusion was no longer possible, one quarter of the early hydrogen (by mass) had been converted to helium. This proportion is essentially the same that we see today.

All elements heavier than iron were not produced in the big bang, but were instead created in the heart of supernova explosions.

1.72 The Fate of the Universe

Fate of Universe

Ever since astronomers have realized that the universe was expanding, they have been trying to determine its ultimate fate. Will it go on expanding forever, or will it eventually all collapse back into a single point? Several key issues must be addressed in any study into the fate of the universe.

A. MASS DEPENDENCE

The fate of the universe depends to a large extent upon its total mass. Is there enough mass (and hence gravity) to halt the expansion and recollapse the universe? If there is, we say that the universe is "closed". If there is not enough matter, so the universe keeps expanding forever, we say that the universe is "open". The mass density of the universe is usually specified by the greek letter Omega. If Omega<1, the universe is open, and if Omega>1, the universe is closed. Astronomers are currently performing a multitude of experiments and observations to help determine the mass of the universe, and ultimately, Omega.

B. CRITICAL DENSITY

The critical density is the boundary case between the scenarios where the universe has enough mass/volume to close the universe and too little mass/volume to stop the expansion. A formula for critical density may be given by:

$$\text{critical density} = \frac{3*H^{2}}{8*Pi*G}$$

where G is the gravitation constant (6.67e-11 m^3/kg/s^2), and H is the Hubble constant. With the current estimate of the Hubble constant, we can estimate the critical density to be approximately 8e-30 g/cm^3. If we are able to determine the actual density of the universe, we can then calculate Omega by dividing the current density by the critical density. Unfortunately, a determination of the current density is proving very difficult to do.

C. CHANGING HUBBLE CONSTANT

Despite its misleading name, the Hubble constant has in fact been different at different cosmological times. Since gravity slows expansion, the early universe was expanding faster than it is now. This means that the critical density was greater at earlier times. However, the changing Hubble constant doesn't imply that the universe can change its status from open to closed. If a universe starts out with a density greater than its critical density, then this will always be so (and vice versa). This is because the density of an expanding universe is always decreasing as its critical density decreases.

D. COSMIC INVENTORY

So what happens if we try and account for all the known matter in the universe in an effort to determine if the universe is open or closed? It turns out that we have 10 to 20 times too little mass to close the universe. However, astronomers have recently detected evidence of "dark matter" - matter which we cannot see but whose gravitational effects can be felt. Since some people estimate that there may be more than 10 times the amount of dark matter as visible matter, the universe may in fact be closed. More data and research is clearly needed to determine the ultimate fate of the universe.

1.73 Requirements for Extra-Terrestrial Life

Much of the following chapter is based on information provided by Nick Strobel. It has been edited, and included with "The Digital Universe" with the kind permission of the author.

SOME REQUIREMENTS FOR EXTRA-TERRESTRIAL LIFE

* Life Zones and Suitable Stars for E.T. life
* Habitable Planets
* Life Characteristics
* What are the chances of E.T. life? - the Drake Equation

1.74 Life Zones and Suitable Stars for E.T. life

Life Zones and Suitable Stars for E.T. life

To get a feel for the environment necessary for water-based life to develop, we define the "life zone" as the distance from the star where the temperature is between 0 - 100 degrees C. The inner and outer bounds for the life zone of the Sun (a G2 main sequence star) are 0.7 and 1.5 A.U., respectively. To determine the inner and outer bounds of the lifezone for a star like Vega (A0 main sequence star) and for a star like Kapteyn's star (M0 main sequence star), we

```
need the following relations:

 a) luminosity of a star = 4 pi r^2 * sigma T^4 , where r is the
    distance from the center of the star, T is the temperature on the
    surface of the sphere of radius r and pi & sigma are constants.
 b) L_{Vega}/L_{Sun} = 53 and L_{Kapt}/L_{Sun} = 0.004 where L_{star}
    represents the luminosity of a star.
At the outer boundary of the lifezone the temperature is 0 degrees C and the
inner boundary is at 100 degrees C. So:

       L_{star}/L_{Sun} = r^2_{star}/r^2_{Sun}

since the 4 pi and sigma T^4 cancel out. Now we can solve for the distance from
the star (r_{star}) using the lifezone boundaries of the Sun (r_{Sun}= 0.7 &
1.5 A.U.). For Vega we get life zone boundaries of 5.1-10.9 A.U. and for
Kapteyn's star we get 0.044-0.095 A.U.

Now, let's assume that it takes 3 billion years for intelligence to evolve on a
planet.   First consider the lifetime of a star. We'll want the star to last at
least 3 billion years!  This sets the upper limit on our mass to be 1.5 times
that of the Sun (heavier stars would not last long enough).  Lighter stars have
longer lifetimes, but the life zones get narrower and closer to the star.  If a
planet is too close to a star, it will get its rotation tidally locked so that
one side of the planet always faces the star.  Assuming that this is
undesirable for life to develop, the life zone considerations limit the stellar
mass at being 0.5 times that of the Sun.  As a result, we assume that a star
should be between 0.5 and 1.5 times as massive of the Sun if life is to have a
chance at developing.
```

1.75 Habitable Planets

```
For life to develop on a planet, it is not enough that the planet lie within
it's star's life zone.  The planet itself should have:

      * a stable temperature regime
      * a liquid mileau to mix
      * the essential building block elements (carbon, hydrogen,
        nitrogen, oxygen, phosphorus, sulfur, and transition metals like
        iron, chromium, and nickel). Since the building block elements are
        only created in the stars, the best places to look for life is
        around stars formed from processed gas, ie., look at metal-rich
        stars.
      * a solid surface to concentrate the building block elements together
        in the liquid on top.
      * enough gravity to retain an atmosphere.
```

1.76 Life Characteristics

```
If we were to ever stumble across an extra-terrestrial life form, it would
likely be quite different in appearance than ourselves.  How would we know that
it is alive?  From a biology textbook, we can list the following general
characteristics.
```

1. Organization. All living things are organized and structured at the molecular, cellular, tissue, organ, system, and individual level. Organization also exists at levels beyond the individual, such as populations, communities, and ecosystems.

2. Maintenance/Metabolism. To overcome entropy, living things use energy to maintain homeostasis (maintain its sameness; a constant, structured internal environment). Metabolism is a collective term to describe the chemical and physical reactions that result in life.

3. Growth. Living things grow. The size and shape of an individual are determined by its genetic makeup and by the environment.

4. Response to Stimuli. Living things react to information that comes from outside or inside themselves.

5. Reproduction. Individuals reproduce themselves. Life also reproduces itself at the subcellular and cellular levels. In some instances, genetic information is altered. These mutations and genetic recombinations give rise to variations in a species.

6. Variation. Living things are varied because of mutation and genetic recombinations. Variations may affect an individual's appearance or chemical makeup and many genetic variations are passed from one generation to the next.

7. Adaption. Living things adapt to changes in the their environment.

Items 2 and 3 are related. Life grows by creating more and more order. Since entropy is decreased, life requires energy input. Life gains local structure at the expense of seemingly chaotic surroundings on a large scale. Items 5, 6, and 7 are related. Life reproduces - complex structures reproduce themselves. Life changes itself in response to natural selection on the macroscopic level and to changes in DNA on the microscopic level.

1.77 What are the Chances of Extra-Terrestrial Life?

What are the Chances of Extra-Terrestrial Life?

For countless centuries, people have tried to estimate what the chances of extra-terrestrial life were. Most of the theories were entirely speculative. One of the first astronomers to adopt a scientific approach to the problem was Dr. Frank Drake, while he was working at the National Radio Astronomy Observatory (NRAO) in Green Bank, West Virginia in 1961. He presented an equation which estimated the probability of intelligent life. It became universally known as the "Drake Equation", and it is presented below:

```
            N = Rs x Fp x Ne x Fl x Fi x Fc x L

where:
  N  = the number of civilizations in our galaxy whose radio emissions
       are detectable and are theoretically able to communicate with us.
  Rs = the rate of formation of stars with a large enough life zone
```

```
      and a long enough lifetime to be suitable for the development of
      intelligent life.
  Fp = the fraction of stars which are accompanied by planets.
  Ne = the number of planets per star that are ecologically suitable
      for life.
  Fl = the fraction of those planets where life develops
  Fi = the fraction of the above planets where intelligence evolves
  Fc = the fraction of planets with intelligent life that develop a
      technological civilization.
  L  = the length of time that such civilizations release detectable
      signals into space.
```

This equation helps to clarify the factors which must be considered in estimating the probability of intelligent extra-terrestrial life.

Unfortunately, the results of the equation are only as good as our estimates for the various factors used in evaluating it. In the past, results have varied wildly. Some optimistic astronomers believe that one in a thousand stars are the home to an intelligent civilization that we have the potential to communicate with. Other, more pessimistic individuals have used factors which yield probabilities of less than one in a billion. Frank Drake himself estimates that one technological civilization can be found for every 10,000 stars.

Nobody doubts the validity of the Drake Equation. But clearly, much work remains to be done to provide an accurate determination of the various factors which are involved.