Lectures on the History of Astronomy

Version 0.0

David J. Jeffery
Department of Physics and Astronomy
And Dyer Observatory
Vanderbilt University

1996 August 27

♠

♠ ♠ ♠ ♠

♠ ♠

♠

♠ ♠

“from sources close to or out of high school years,
perhaps at parties or in sleep,
word has passed from mouth to street
of something willow weird and stalking mad,
and well, so is it true, is every day recalled,
does neither milk nor sparrow fall unmourned,
are all present, a secret chorus to our speech?”

—Je59

Contents

Preface

Boldly or brazenly, I have written a book of lectures on the history of astronomy. The lectures were prepared for the one-semester history of astronomy course (Astronomy 130) at Vanderbilt University. This course is designated as a Science and the World course. Such courses are intended "to introduce students to scientific and technical knowledge and to relate that knowledge to the broader context of the world." General histories of astronomy such as the fine recent book by John North and the classic book by Anton Pannekoek are too narrowly focused on astronomy to satisfy the intention of the course. They are also often too technical for the nonscience students who are in the majority in my classes.

In order to carry out the course mandate, I have attempted to analyze the historical development of astronomy both in terms of internal intellectual development and in terms of external factors affecting astronomy. Not being a trained historian nor an expert in the details of the history of astronomy itself, I have been forced to rely almost entirely on secondary authors. Sometimes I have sought the best ones available, but often the pressure of time led me to use those sources that were easily available. I hope to seek out authoritative sources for all topics in future versions.

Given the course mandate it is not desirable or possible to cover all of the history of astronomy. I have focused on episodes which I think have the greatest value for understanding the intellectual development of astronomy from earliest times to the present. Clearly, my choice of topics is idiosyncratic. I have also allowed myself some digressions where topics caught my interest.

Experience has shown that nonscience students are generally not at ease in the Zion of mathematical astronomy. I have accommodated this unease to a degree. However, some basic trigonometry and formulae are essential for comprehending the issues. Therefore I have provided some of this material in Chapter 2 and in other places throughout the lectures; some more advanced material appears in Appendix A.

I have chosen an astronomical way of citing sources. A reference to a work in the text is usually made with the first few letters of the author's (or first author's) name followed by the page number if appropriate. The abbreviation for North's history of astronomy book for example is No. In the citation scheme, a reference to page 264 of North's book becomes No264. The citation scheme allows one to go right from the line of text to the page of the source without the bother of ibid.'s and op. cit.'s. The drawback is that digressions on the source material must be abandoned, put into the text, or footnoted. If a work is only cited one or two times, I cite author, date, and, if appropriate, page number in text.

It should be noted that many cases I cite a source only as an example reference. Since this is so common, I have refrained from using 'e.g.' (*exempli gratia*: for example) in most cases. As well as 'e.g.,' other common abbreviations that I use are 'c.' (circa), 'd.' (died), 'esp.' (especially), 'ff.' (and the following), 'fl.' (flourished), 'i.e.' (*id est*: that is), 'r.' (reigned), and 'viz.' (*videlicet*: that is to say; namely),

The full reference for any work appears in the references section. The abbreviation for the work is in parentheses at the end of the full reference. I should mention that North (No624ff) provides a good bibliographic essay with references to many other sources for the history of astronomy. A more extensive (alas now slightly dated) selected, annotated bibliography of the history of astronomy from the invention of the telescope is given by DeVorkin (1982).

This is version 0.0 of the lectures (with typesetting in plain TEX [Knuth 1986]). It is fragmentary and unpolished. Multiple question marks at the end of a word or statement indicate that I am not sure of the spelling (??), reference (???), or fact/assertion (????). It must be understood that almost all references are to secondary sources and that almost all of these are quoted as example references for facts, ideas, or analysis, not as original ones (although they may be that too). There could be inconsistencies, redundancies, unclearness, and mistakes: *caveat lector*. I would appreciate any suggestions or learning of any errors.

I would like to thank my students and colleagues at Vanderbilt for their encouragement and aid.

David J. Jeffery
Vanderbilt University
August 27 1996

1. Introduction: Astronomy, History, and Science

1.1 Astronomy and History

Astronomy comprehends the universe: that is to say, it studies the universe as whole, things of largest scale, things of smallest scale, and everything in between is touched, and it watches the sky. Fundamental physics is its serf, its master, its synonym whenever it, astronomy, is the One and not the Many. At astronomy's boundaries—and no one marks those with a stone—are all physics, geology, planetology, biology, exobiology, astrology, philosophy, the human condition, society, and—rightly? wrongly?—religion. Astronomy is history: the story of our peaceful studies going back to proverbial time immemorial and the story of the evolution of everything. Both stories proceed. In brief, whenever you lift up your head you find astronomy, find it is always all around you, and find you are part of it.

Since Plato (more or less), the supremos of pedagogy have placed astronomy on a pedestal: made it one of the seven liberal arts "which liberate the soul from its native darkness and ignorance" (Pe166). The planners of modern liberal education have swept it off that eminence; they have determined it is only one science among many. But even in this humbled category, astronomy has a faded glory. Its broad scope and deep anchoring are attested by its popularity: the people have always favored it. It is the first empirical exact science (Noxiii) and is probably coeval with with mathematics. Indeed, until the end of the 18th century leading theoretical astronomers and mathematicians were often the same people. As the first empirical science, astronomy has naturally led—often albeit without any followers. The method of science—hypothesis, verification, nonverification, new hypothesis on a wider plane—was early on most clearly exampled in astronomy. And once, in the 16th and 17th centuries, it took the starring role in that transformation, the Scientific Revolution, which took science from occasional lofty peaks and long lethargies to continual progress. This same continual scientific progress, feeding and being fed on technology, has circled the globe, transforming everywhere and then transforming again. At present, we do not foresee an ending to the transformations; indeed we do not want them to stop, only to direct them—if only we knew exactly where.

To connect astronomy to the human world we have, among other resources, the history of astronomy which is the subject of these lectures notes. This history comprises the purely intellectual give and take of theories, the personal histories of the astronomers, and the interaction of astronomy and astronomers with society. The first element can be viewed, bloodlessly, as an autonomous internal development. This has the advantages of conciseness and retrospective plausibility; it is useful for the fast overview and for the establishing the basis for later astronomical achievements. However, any real history shows that astronomers did not proceed so simply: there was no path through the thicket. The astronomers lived in a broad intellectual world that was a congery of different, not always compatible, ideas: some profound, some the opposite; some are still with us, some we have done without a good long while. Especially for older times this blooded, broad intellectual world cannot be overlooked without missing much; and we will not overlook it, not entirely at least. In more recent times, the 19th and 20th centuries, the broad intellectual world of scientists has become more standardized and familiar: sceptical, accepting partial answers as good answers, and based on the bedrock of the modern scientific method. Having seen this intellectual world view develop in our course, we will have less to say about it when it is in full bloom in the now.

The other elements, the personal histories of the astronomers and the interaction of astronomy and astronomers with society is in our purview. In fact, these elements cannot be cleanly separated from the intellectual element. The astronomer does not abstractly absorb his/her ideas from their times and work on them without regard for the wind and weather. Each one somehow selects a subset of current ideas and weights them individually. Then somehow amid distractions—everything that is not astronomy—the astronomer does their work and achieves or does not achieve anything. How the astronomer was educated, earned a living, and was applauded or reviled are relevent to their achievements. In truth, the truths (speaking as we speak) that they learn are indifferent to the seeking, but we are not. How, human achievement is achieved is among our studies.

How astronomy interacts with society is a fundamental part of these lecture notes; we will not do it justice, but we will do what we can. We will mention time keeping. Day and night follow day and night. The years proceed. Though we little note it now, the Moon waxes and wanes just as formerly it did. The planets are esoteric, but remain regular too. Digital watches and atomic clocks notwithstanding, astronomy still marks our days and works.

We will mention also astrology. Deride if you will (and we will deride it), for most of historical astronomy, astrology was a key practical benefit in the eyes of society. And many astronomers before 1650 or so earnt their living, gladly or not, prognosticating.

Then there is the secret of the universe, life, and everything. The yearning for this secret gnaws somewhere. At times and often even, the secret has been declared revealed. But a general acknowledgment has been withheld. Some have always thought that astronomy was one of the most promising avenues to the secret. Modern astronomers, with the modern scientific orthodoxy of partial answers, have generally ducked metaphysics. However, when the subject is cosmology (everything and itself), it cannot be denied that a meeting up somewhere with metaphysics is possible. In fact, the venture might seem lastly profitless if the meeting is eternally postponed.

The search for extraterrestrial intelligence (SETI) also connects astronomy and society. A band of astronomers and an element of society think SETI of immense importance. Contact may well change our way of thinking in ways we cannot now think of. Given that success would be shocking and failure inconclusive, it is likely that we will see the search continue all our lives. Since our capacities to communicate to the stars will likely increase in time, SETI will probably loom ever larger in the future history of astronomy.

The main path that we will follow through the history of astronomy is one that a modern astronomer would identify as a progressive path. This path will take us through the astronomies of the prehistory (but only briefly), the ancient Mesopotamians, the ancient Greeks, the Medieval Islamic civilization, Europe and America from the Renaissance through the 19th century, and the modern age. Along this path a continual advance in astronomy (as seen in the telescoping view of history) will be apparent. For a survey course, this is pedagogically a sound. Moreover, modern astronomy and the path leading to it is generally most interesting to us moderns.

Given the underlying theme of the lecture notes—science and the human world—it would naturally be valid to study those astronomies that did not directly lead to or contribute much to modern astronomy: e.g., those of ancient Egypt, China, Japan, India, and pre-Columbian America. However, we do want to exhaust ourselves and lose a clear vision in an encyclopedia of astronomies. Therefore, we will move off our main path only occasionally.

1.2 History and Science

The history of astronomy is, of course, part of the history of science and of history as a whole. If one entirely isolates the history astronomy from rest of history, one is forced to avoid important interactions and to deal in explanations of astronomical development that must be incomplete. It is well beyond the scope of these lectures notes to treat the complete set of relations between astronomy and the rest of the human experience. But, as mentioned in § 1.1, we attempt to go some distance in that direction, particularly for the pre-modern and early modern period. The relation of science and society in these times is of importance because we are tremendously interested in how the transformational powers and intellectual adventures lurking in science came to be revealed and unleashed.

We need to be acutely aware of the limitations of historical explanation when analyzing events. One famous historian of science, Otto Neugebauer, has the following to say:

> This is a good illustration for the futility of any attempt to reconstruct "reasons" for the incidents of historical events. Similarly the absence of algebraic notations should not

> have prevented the Greek geometers from developing what was called in the 19th century "synthetic" and "projective" geometry since many of the basic concepts were ready at hand in the works of Apollonius. Again such a "natural" development did not take place and all that we may ever hope to establish in historical research is facts and conditions but never causes (Ne225).

I think a key word in this quotation is "establish." We cannot establish historical causes as an empirical exact science can physical causes. But causes surely exist, and I do not think that Neugebauer can mean that they are in principle unknowable. Another quote from an eminent historian of science and technology, Lynn White, hits closer to the mark I think:

> Historical explanation . . . is seldom a matter of one billiard ball striking another, of 'causes' in the narrow sense. It is much more often a process of gradual illumination of the fact to be explained by gathering around it other facts that, like lamps, seem to throw light on it. At last the historian arrives at a sense that the central fact on which he is focusing has become intelligible.
>
> [This is] the sort of explanation, necessarily common among historians dealing with large phenomena, that can neither be proved nor disproved with any rigor but that must be accepted or rejected on grounds of general coherence or incoherence (White 1978, p. 217, 333, taken from Co517).

No gloss is needed.

Much of the historical analysis of the development of science given in these notes is based on the recent book *The Scientific Revolution: A Historiographical Inquiry* by H. Floris Cohen (Co). This book is a analysis of the understanding and explanations for the Scientific Revolution, a term used for the development of early modern science in the 16th and 17th century in Europe. The term 'early modern science' for the science of the 16th–17th centuries is Cohen's usage (Co17). The Scientific Revolution is a unique event. However, the term 'Scientific Revolution' is also used for hypothetical events giving rise to science like modern science that might have occurred in other times and places. Scientific Revolution is always capitalized in order to distinguish Scientific Revolutions from the scientific revolutions in the model of scientific development of Thomas S. Kuhn in his famous book *The Structure of the Scientific Revolutions* (1962). Kuhn's small-s, small-r scientific revolutions are different beasts; they are relevant also to the history of astronomy, but we will not discuss them explicitly.

Modern science has two defining characteristics: it based on the modern scientific method and it has turned out to be continuously progressive and cumulative. The modern scientific method consists, in brief, of a continuous cycle of theory, experiment or detailed observation, leading to improved, generalized theory. Although we loosely speak of a theory as being true or false, what we mean is adequate or false. An adequate theory accounts for what we observe: modern science claims no access to absolute truth. Nevertheless, some theories are so well established that we describe them as true when we are not speaking with absolute precision. Theories such as Newtonian gravity (within its range of known applicability) are 'true' in this sense. It is generally believed that science, however, is on an approach to absolute truth about the natural world; perhaps it is a limit that can never be reached or the idea of continuous approach is a delusion; but the belief that we are on the approach is the faith.

A key aspect of modern science, always present implicitly in it, but strongly emphasized by the 20th century philosopher Karl Popper, is that a scientific theory must be falsifiable. A theory not falsifiable in practice is suspect; one not falsifiable in principle is clearly not scientific. Unfalsifiable-in-principle theories may be true, but their elucidation and defence are not in the realm of science.

The continuously progressive and cumulative feature of modern science which started in Europe and has spread world-wide is a historical fact. It is based on human society as it exists and on the continuing comprehensibility of reality. Neither bases need persist forever, but at present they seem securely in place. The progress and accumulation of science have changed the intellectual and material world over and over again. They will probably continue to do so. Thus science is one of the overwhelmingly important elements of our lives. It is also very unpredictable. There is an ancient theory that history is cyclic; this theory is evidently false in the world of modern science.

But the world of modern science did not exist before about the 16th century. Science in earlier times can probably judged to be cumulative overall, but it was not continuously progressive: it was sometimes progressive, sometimes static, sometimes in regression, and sometimes non-existent. All modes could be happening simultaneously in different parts of the world. In the model of scientific progress of Joseph Ben-David (discussed by Co254ff), these varying modes of science are natural: modern science is exceptional. Societies in which science is not modern are called traditional societies. In traditional societies, occasionally progressive science is the best that occurs. The explanation for the weakness of science in traditional societies requires a great deal of analysis. But two causes can be plausibly cited: one external and one internal. First, the total support for science was weak because society in general did not recognize science as a key concern. Second, the scientific method was not clearly recognized, and so what we think of as the best way to do science was not always used. We will interpret the particular cases of ancient Greek science (§ 4.12) and Islamic science (§ 5.7) using the Ben-David model. A general discussion of the Ben-David model is beyond our scope.

Some terminology that we will use when using the Ben-David model should be introduced here. We distinguish science in traditional societies as dormant, sporadic, or golden-age. Dormant means a virtual absence of science for a prolonged period of time. An example of dormantness is science in the Byzantine empire from circa 600 to a very minor re-awakening in the 14th centuries (Pe151; No223). During this period the science of ancient Greece continued to exist in manuscripts copied and recopied, preserved, summarized, and commented on, but there were hardly any innovations and no significant individual scientists or natural philosophers of note. Some dormant societies, of course, were never awake: most pre-literate societies are examples.

Sporadic science means that here and there an individual or a small group of innovators appears. But they are largely isolated geographically and temporally. However, if we have heard of them at all, some or their works survive and may have had a historical effect.

Golden-age science is not entirely a neologism, but my usage needs some specification. I use it to describe scientific development that is continuous for a prolonged period of time, but that ultimately stagnates. Greek science in the period circa 600–200 BC was golden-age: this is my prime example (see § 4.12). The period 800–1000 AD has been identified the Golden Age of Islamic science by Sayili (1960). This may be another golden-age in my sense, but I am not sufficiently informed about the Islamic science to make a judgment. The term golden-age science may have limited value since one person's golden age is another's vigorous sporadic science.

The terms decline, stagnation, and fading describe the transition from sporadic science to dormant science or golden-age science to sporadic science. This usage seems general in historiography (e.g., Co250ff, 384ff, and references therein).

2. Stars Above: The Basics of Astronomy

There is no royal road to astronomy.[1] As in any any other study, one has to master basic terminology, concepts, and problems.

2.1 Angular Measure

The sky has no readily apparent depth. All celestial objects seem to be pasted on a dome, or rather a sphere that rotates around the Earth daily. This fictitious, but useful, sphere is called the Celestial sphere. (The capitalization in the term Celestial sphere is necessary to distinguish it from the historically important celestial spheres introduced by the Greeks.) To this day, measuring the distances to objects in space is one of the hardest and most controversial of astronomical measurements: of course, we have the easier objects measured now. Angular measurement of position on the Celestial sphere, however, is relatively simple. Here we present some of the basics of angular measurement.

The circle is traditionally divided into 360°, where superscript circle stands for degrees. Degrees are the basic units of angular measure. Although many symbols are used to represent an angle, the most common is the Greek theta: θ.

The angle between two objects is their angular separation. It is common to say that the angle subtends the line or arc between the two objects (Ba1207). From any point on the Earth, the angular separation between astronomical objects is almost always the same to such a high accuracy that the difference due to different vantage points is negligible. Such a difference is called a parallax.

Crude, but useful, angular measurements can be done with the human hand held at arm's length. The angle separation across a finger, or to use more common words, the angle subtended by a finger is $\sim 1°$ (Ze5). The angles subtended by a fist and a spread hand are $\sim 10°$ and $\sim 18°$. As an example, the angle between the pointer stars of the Big Dipper (the stars approximately aligned with Polaris: see also § 2.2) is about the angle subtended by half a fist at arm's length: i.e., $\sim 5°$ (Ze5).

By ancient tradition degrees are not divided into decimal fractions, but into sexagesimal fractions. One degree is divided into 60 arcminutes and 1 arcminute is divided into 60 arcseconds. The traditional symbol for arcminute is the prime symbol and for arcseconds the double prime symbol. To summarize,

$$1° = 60' = 3600'' \tag{2.1}$$

and

$$1' = 60'' \ . \tag{2.2}$$

The smallest angular separation that can be resolved by the naked eye is $\sim 1'$ (Ze49). However, systematic naked-eye measurements almost never achieved this accuracy. The last and greatest pre-telescopic astronomer, Tycho Brahe (1546–1601) (Th4, 469), was achieving better than $1'$ accuracy in later work, but this was unprecedented (No302). Because the telescope was invented shortly after

[1] The royal road metaphor seems to have originated with Euclid. Proclos of Lycia???? (c. 412–485 [Pe382]) reports the following anecdote:

> Ptolemy once asked Euclid whether there was any shorter way to knowledge of geometry than by study of the Elements, whereupon Euclid answered that there was no royal road to geometry (quoted from Bo111).

The Ptolemy in the quote is not the astronomer, but Ptolemy I????, the first Macedonian king of Egypt and formerly one of Alexander the Great's generals.

his death, Brahe's naked eye work has probably never been bettered at least on a large scale, except perhaps by the distinctly old-fashioned master observer Johannes Hevelius (1611–1687) (Pa259; No342).

The angle subtending the line bisecting the disk face of a spherical astronomical object is called the angular diameter. The only two common spherical astronomical objects for which shape can be resolved with the naked eye are the Sun and the Moon. The mean angular diameters of the Sun and Moon are $32'$ (Mo242) and $31'5'' \approx 31.1'$ (Mo93), respectively. The near equality of the Sun and Moon's apparent size is just a coincidence: part of the geography of the solar system. But to early peoples, to whom the Earth was not just a planet, this near equality seemed fraught with divine significance. The Greeks, for instance, considered the gods of the Sun and Moon to be twins, Apollo and Artemis.

The division of the circle into $360°$ degrees is was an ancient convenience. For the users of a decimal numeral system it is an inconvenience that seems ineradicable. There is, however, another division into radians. This division is natural: i.e., suggested by the nature of Euclidean geometry rather than based on human convention. The angle in radians between any two points on a spherical shell is the circular arc length between the points divided by the radius of the arc. If arc length is s and radius is r, the angle in radians is given by

$$\theta = \frac{s}{r} \ , \qquad (2.3)$$

where s and r are measured in the same units. Since the circumference of circle, C, is given by

$$C = 2\pi r \ , \qquad (2.4)$$

if follows that the conversions between radians and degrees are given by

$$\theta(°) = \frac{180°}{\pi}\theta(\text{radians}) \qquad (2.5)$$

and

$$\theta(\text{radians}) = \frac{\pi}{180°}\theta(°) \ . \qquad (2.6)$$

Although it is not especially useful to know this,

$$1\,\text{radian} \approx 57.2958° \approx 57°17'45'' \qquad (2.7)$$

and

$$1° \approx 0.0174533\,\text{radian} \ . \qquad (2.8)$$

A very crude approximation that I find easy to remember is

$$1'' \sim \frac{\pi}{60^3}\,\text{radian} \ . \qquad (2.9)$$

There are several reasons for preferring radians to degrees. First, in using radians if any two of arc length, distance, and angle in radians are known (with the distance in the same units), then the third quantity can be determined without requiring conversions. More importantly, in algebraic work using radian measure avoids the need for inelegant conversion constants to appear in formulae. Most importantly, calculus formulae involving angles are written naturally and simply using radian measure.

The final topic of this section is angular velocity, the rate at which an object changes its angular position with the angle being measured from a fixed point. One common convention is to represent angular velocity by $\dot{\theta}$. This dot notation goes back to Newton in 1671 although not made public until much later; he used the dot to indicate a fluxion which we would call a derivative (Bo435; Ha253). Angular velocity can be measured in degrees, arcminutes, arcseconds, radians, or other angular units (see, e.g., § 2.2) per unit time.

If an object goes in a circular orbit about a central point with a constant spatial velocity, then its angular velocity in degrees per unit time is

$$\dot{\theta} = \frac{360^\circ}{P} \ , \tag{2.10}$$

where P is the period of the orbit. In general, angular velocity will not be constant and the spatial motion will not be circular about a point. Nevertheless, the mean angular velocity about any point by an object whose motion completely surrounds the point will be given by equation (2.10) with $\dot{\theta}$ replaced by $\dot{\theta}_{\rm av}$. If an object traverses an angle $\Delta\theta$ (as measured from some point) in a time Δt, then its mean angular velocity during Δt is given by

$$\dot{\theta}_{\rm av} = \frac{\Delta\theta}{\Delta t} \ . \tag{2.11}$$

When Δt becomes very small $\dot{\theta}_{\rm av}$ approaches the instaneous velocity $\dot{\theta}$ which is the derivative of θ with respect to time in the terminology of calculus.

2.2 The Horizon Coordinate System

Several angular coordinate systems are used in astronomy to locate objects on the Celestial sphere. The two basic ones are the horizon system and celestial (or equatorial) system.

The horizon system is a local system: i.e., a system dependent on one location on Earth. The origin or center of the system is where you are. The direction of gravity (determined by a plumb line) is the normal (or perpendicular) to the horizon plane. Only if one were on a very flat piece of ground would the horizon plane be truly tangent to the ground. However, crudely speaking one can see the sky above the horizon and not below. Directly over head is zenith (from the Arabic *semt ar-arās* meaning way over head [Ba1418]). Directly opposite to the zenith (and so below the Earth) is the nadir (from the Arabic *nazīr* meaning corresponding or opposite [Ba806]). The four cardinal points of the compass are due north, east, south, and west points.

The polar axis lies in the plane defined by the north and south points, and the zenith. The polar axis is generally titled with respect to the horizon plane. A bit of geometry shows that the angle of the polar axis from the north south line is equal to the latitude of one's position. Because astronomical distance are so huge compared to the radius of the Earth, for the most part one can think of oneself has being at the center of the Earth position on a horizon plane that is titled from the pole by the local latitude. In the northern hemisphere, the north pole is above the horizon and the south pole is below; the reverse situation holds in the southern hemisphere. In the northern hemisphere, the Polaris (also the Pole Star and α Ursa Majoris???) is can be used to easily identify the north pole at night. Polaris is fairly bright and is $44'9''$ (in epoch 2000 measure) from the true pole (Ho20). Polaris itself can be easily identified from the two pointer stars of the Big Dipper. These stars form the bowl edge farthest from the handle. Going along the line marked by the pointer stars in the direction above the bowl leads to Polaris.

Two angles are needed to locate an object on the Celestial sphere: the azimuth and the altitude. The great circle from the zenith through the object cuts the horizon in a particular direction. The angle to this direction measured westward from due south is the azimuth (Mo15). The angle from the particular direction to the object is the altitude. One has to note that astronomical altitude is an angle not a height above the ground or sea level.

Now the Earth rotates once per day on its axis. But in the horizon system this means that the Celestial sphere rotates once per day the polar axis turning westward. The stars, Sun, Moon, planets, and other celestial objects are carried about by the Celestial sphere. For short periods of time these objects seem to be fixed in relative positioni although their azimuths and altitudes change continuously. The stars in fact hold their relative position for times long compared to most ordinary human time scales. Thus we talk of the fixed stars. The Sun, Moon, and planets, in fact move noticeably relative to the fixed stars on ordinary human time scales. There basic motion is eastward relative to the fixed stars. The Sun and Moon always move eastward. The planets move eastward most of the time. When the planets move westard relative to the fixed stars they are in retrograde motion.

Many celestial bodies rise in the east and set in the west. In the northern hemisphere, those objects that are sufficiently near the north pole never rise or set: they are always above the horizon. Such objects are called circumpolar. Objects too near the south pole are never above the horizon. The situation is reversed in the southern hemisphere.

The great circle or meridian on the Celestial sphere running through the poles is the celestial meridian (Mo15) or commonly just the meridian (as opposed to any old meridian). As the Celestial sphere spins westward all celestial objects cross the meridian twice per day: in the jargon they transit the meridian twice per day. When they transit the part of the meridian that is mostly above the horizon that is called upper culmination; lower culmination is when they transit the part of the meridian that is mostly below the horizon. Usually upper culmination is meant when the expression transiting the meridian is used. Only circumpolar objects have both upper and lower culminations above the horizon. Objects that are too close to the pole below the horizon have both culminations below the horizon of course. The upper and lower culminations of the Sun have special names: noon and midnight. The astronomical noon and midnight are only approximated by our conventional clock noon and midnight.

The spinning of the Celestial sphere means that in the course of a day objects trace out circle around the pole: the pole is exactly perpendicular to the plane of the circle. The circles are naturally tilted, not perpendicular or parallel to the horizon, except at the Earth's equator and poles, respectively. At upper culmination the object reaches its highest altitude in the day. Observing the highest altitude is a way to identify upper culmination and and the north-south line.

The horizon coordinate system has limited usefulness because objects are always changing their locations. A system in which the fixed stars are at fixed positions is needed. This is provided by the celestial coordinate system. In simple practical astronomy, one often needs to mentally juggle back and forth between the horizon and celestial systems in order to best understand what one is seeing.

2.3 The Celestial Coordinate System

The celestial coordinate system is analogous to the ordinary geographic longitude and latitude coordinate system. To construct the celestial coordinate system project from the center of the Earth the north and south poles, and the equator on the Celestial sphere; these projections become the north and south celestial poles and the celestial equator. We now set up an angular coordinate system on the celestial sphere that is reasonably fixed with respect to the fixed stars: it is not perfectly fixed for reasons we describe below. Naturally this coordinate system rotates every day, but that overall motion is easily accounted for.

In the celestial system, as in the horizon system, two angles are needed to specify any point One angle is measured along a meridian from the celestial equator. This angle, which is analogous to latitude, is called declination and is traditionally represented by the small Greek delta, δ. There are north and south declinations. Declination is measured in degrees, arcseconds, and arcminutes.

The angle corresponding to longitude in the celestial system is right ascension abbreviated to R.A. The zero meridian of right ascension passes through the spring equinox. The spring equinox which will explain below is a specific point on the celestial equator which is located in the constellation Pisces. The right ascension angles are measured eastward from the zero meridian or, in the jargon, eastward from the spring equinox. Right ascension is not measured in the ordinary angular units, but in special units called hours, minutes, and seconds. The conversions to the ordinary angular units are:

$$1\,\text{hour} = 15^\circ \ , \tag{2.12}$$

$$1\,\text{minute} = \frac{1}{4}^\circ = 15' \ , \tag{2.13}$$

and

$$1\,\text{second} = \frac{1}{4}' = 15'' \ . \tag{2.14}$$

The rationale for these peculiar units is simple: the Celestial sphere turns one hour, minute, or second in angle in the corresponding of time. Thus, any angular distance in right ascension units is the same as the time it takes the Celestial sphere to turn that distance. This is a useful correspondence. For example, if the spring equinox transits the meridian in one's location at time $t_{\rm v}$, then the time that object X transits the meridian, t, is given by

$$t = t_{\rm v} + {\rm R.A.(X)} \ , \tag{2.15}$$

where R.A.(X) is the right ascension of object X.

It can be remarked that if angular and time measure were reduced to a consistent simple decimal system, much of the confusion of the inconsistent archaic system we now use would go away. But we seem to be stuck with the archaic system. As one's parents always say "if it was good enough for the Sumerians, it's good enough for us."

The Sun, Moon, and planets move in the celestial system. These motions are easily understood in the modern heliocentric picture of the planetary system. The Earth and planets orbit the Sun with all the orbits nearly in the same plane; the Moon orbits the Earth in almost the same plane. Thus in the celestial system the Sun, Moon, and planets travel on nearly coincident great circle paths on the celestial sphere.

As is well known, the polar axis of the Earth is tilted by about 23.5° degrees from the normal to the plane of the Earth's orbit around the Sun. Thus, the great circle path of the Sun, called the ecliptic, is titled by the same amount from the celestial equator. The normal to the plane of the ecliptic (the plane of the Earth's orbit) is called the ecliptic axis. The two points where the ecliptic crosses the celestial equator are the spring and fall equinoxes. The Sun takes one tropical year to cover the ecliptic, and so passes the passes the spring and fall equinoxes once each in a year. These passages, of course, mark the beginning of spring and fall. The Sun moves eastward on the ecliptic; this slow eastward motion is superimposed on the daily westward motion. The Moon and planets also follow the ecliptic path. The Moon always moves eastward. As mentioned in § 2.2, the planets move sometimes retrograde: i.e., move westward.

The celestial coordinate system is not ideal. The most important reason for this is the precession of the equinoxes. We will discuss the precession in detail in § 4.10. Here we will just state that the axis of the Earth and thus the axis of the celestial system precess about the ecliptic axis: i.e., they rotate tracing out a cone with an opening angle of approximately twice 23.5°. The precession rate is 1.3969712° per Julian century (See Table B2 in Appendix B for a note on the Julian calendar.) It takes about 25800 years for a complete precession. If the rate were exactly constant (which it is not quite????) the period of precession would be 25770.037 Jyr.

The precession of the north celestial pole and the celestial equator about the north ecliptic is anticlockwise when looking up to the north ecliptic pole. Thus the equinoxes (the points there the celestial equator cuts the ecliptic) precess westward relative to the fixed stars: hence the expression the precession of the equinoxes. Both the declination (measured from the celestial equator) and right ascension (measured from the spring equinox) of the fixed stars must change. The coordinates for the fixed stars are therefore given for particular epochs, not once for all. Currently, fixed stars are usually located in epoch year 2000 coordinates. For precise work corrections for current date are used.

The complication of the slow variation of the coordinates is compensated for by the ease of accounting for the Earth's daily rotation in the celestial system. Moreover, since the fixed stars are not truly fixed, there can be no system in which the coordinates of the fixed stars are unchanging. Thus there is no ideal system. It might be remarked that the absolute standard of rest is now believed to be established by the average motion of objects on the cosmological scale and by the cosmic background radiation CBR.

There a couple of consequence of the precession of the equinoxes that should be mentioned here. First, because the equinoxes are shifting slowly westward along the ecliptic relative to the fixed stars and the Sun travels eastward along the ecliptic, the time periond that it takes the Sun to go from spring equinox to spring equinox is shorter than the time period it takes for the Sun to return to the same place relative to the fixed stars. The first time period is the tropical year (365.24219878 days) and the second is the sidereal year (365.25636556 days). The tropical year is, of course, the seasonal year. The sidereal year, besides its physical relevance, has importance for understanding the nature of the heliacal rising of stars in ancient historical astronomy. We will discuss the heliacal rising and settings in § 2.8. The second consequence of the precession of the equinoxes is astrological; we discuss this in § 2.11.

2.4 Trigonometry

Trigonometry (plane trigonometry to be precise) and spherical trigonometry are essential tools for the analysis and manipulation of angular measurements. The word trigonometry is Greek and means triangle measurement (Ba1295). Trigonometry is deals with the relations of the sides and angles of triangles. Plane trigonometry deals with triangles on planar surfaces; spherical trigonometry with triangles on spherical surfaces.

Consider a right triangle (one with an angle of 90° degrees. The side subtended by the right angle is the longest side: this side is the hypotenuse. Let the length of this side be h. Let the lengths of the other sides be x and y. Let θ be the angle between the hypotenuse and the side of length x. This definition of θ makes the sides of length x and y the adjacent and opposide respectively. The ratios of the lengths of the sides are functions of the angle θ. The three commonest trigonometric functions are given by

$$\text{opposite} \quad \text{over} \quad \text{hypotenuse} = \frac{y}{h} = \sin\theta \ , \tag{2.16}$$

$$\text{adjacent} \quad \text{over} \quad \text{hypotenuse} = \frac{x}{h} = \cos\theta \ , \tag{2.17}$$

and

$$\text{opposite} \quad \text{over} \quad \text{adjacent} = \frac{y}{x} = \tan\theta = \frac{\sin\theta}{\cos\theta} \ , \tag{2.18}$$

where sin, cos, and tan are the sine, cosine, and tangent functions respectively. Trigonometric functions are transcendental functions because they cannot be evaluated by any finite number of algebraic operations (i.e., addition, subtraction, etc.) (Ba1286). Consequently, the values of the trigonometric functions are tabulated or calculated from some algorithm.

A simple example of an astronomical use of trigonometry is the determination of the altitude of the Sun. The simple measurement of the length of the shadow of a gnomon (a vertical stick) along with the height of the gnomon allows one to determine the tangent of the altitude; from the tangent the altitude itself is obtained. The shortest length of the shadow in a day gives the Sun's maximum altitude; this occurs when the Sun is transiting the meridian.

A key use of trigonometry is in parallax measurements, a topic we take up in § 2.5.

2.5 Parallax

Parallax is the shift in angle position of some object relative to some distance background reference point due to the motion of the observer. Parallax is noticed all the time in everyday life. A person takes a step and nearby objects shift in position relative to distant objects. The more distant the object the smaller the shift. Very distant objects appear to have no shift for small observer motions and are effectively at infinity.

Measurement of parallax and the displacement of the observer allows the distance to the object exhibiting parallax to be determined from trigonometry. Distance measurement from parallax is the most ancient method of distance measurement in astronomy and it is still of fundamental importance. There are two simple parallax cases that arise in astronomy; we will call these the sine and tangent cases.

For both cases consider two points: A and B. The we wish to measure the distance, d, from A to B; we are located close to A. There is also a distant reference object. The reference is so distant that it exhibits no parallax (relative to other distant objects of course) no matter how we move.

Initially, we, the reference, and A and B are all on a line: this is line 1 In the sine case we are a distance r from A in the direction of B, but short of B; in the tangent case, we are at A. In the sine case, we move by rotating around A, maintaining our distance r from A, until the angle we measure between A and B is 90°. In the tangent case we move in a straight line a distance r perpendicular to line 1. In both cases, we, A, and B form a right triangle after having moved: in the sine case we are at the right angle; in the tangent case A is at the right angle. We call the displacement r between us and A the baseline of the parallax measurement.

Initially there was no angular displacement between B and the reference. After we have moved there is an angular displacement θ between B and the new line to the reference, line 2. This angular displacement is the parallax. Line 2 must be parallel to line 1 since the reference exhibits no parallax itself. Because lines 1 and 2 are parallel, θ is also the angle at the B corner of the triangle. Using the trigonometric functions, we find the distance d between A and B to be given by

$$d = r \begin{cases} (\sin\theta)^{-1} & \text{for the sine case;} \\ (\tan\theta)^{-1} & \text{for the tangent case.} \end{cases} \tag{2.19}$$

The significance of the terms sine and tangent case are now clear.

In astronomy the parallaxes are usually extremely small, and thus small argument approximations can be used for the trigonometric functions. Note that the small argument approximations are only valid for angles measured in radians, unless explicit conversion factors are included. Using the small argument formulae found in Table B9 in Appendix B for sine, tangent, and the infinite geomentric series, the expressions for the distance d become

$$d \approx \frac{r}{\theta} \begin{cases} \left(1 + \frac{1}{6}\theta^2\right) & \text{for the sine case;} \\ \left(1 - \frac{1}{3}\theta^2\right) & \text{for the tangent case.} \end{cases} \tag{2.20}$$

These formulae are valid as long as θ^4 can be considered negligible compared to 1 in whatever calculation is being done. The terms containing θ^2 are the called the correction terms. If θ^2 can be considered negligible compared to 1, then the correction terms can be dropped and the sine and tangent case formulae become identical.

In most astronomical parallax measurements $\theta < 1°$ and often $\theta << 1°$. Since $1° \approx 0.174532$ radians, even the correction terms in equation (2.20) will often be negligible; the correction terms are in fact usually smaller than the error in the measurement of θ itself. We will, however, retain the correction terms just to show explicitly how the formulae should be corrected if necessary.

There are two traditional baselines in astronomy: the equatorial radius of the Earth and the mean Earth-Sun distance (which is called the astronomical unit with symbol AU). The first arises from the Earth's daily rotation and the second from the Earth's yearly revolution about the Sun. If a parallax is given for solar system object without a cited baseline, then the equatorial radius is usually the baseline used. For objects beyond the solar system, the astronomical unit is usually the baseline used.

When we use the equatorial radius baseline, we use the sine case formulae. This is because we want the distance from the Earth's center to the object exhibiting the parallax in general and because as the object sets on our horizon the angle between the Earth's center the object is 90°. As an example, we can imagine measuring the Moon's parallax. Say we were on the Earth's equator and saw the Moon transiting the meridian. We note that it is aligned with a distant star. Later the Moon sets at which time the angle between the Earth's center and the Moon is 90°. The angle between the Moon and the distant star is the parallax from which the distance can be determined using the sine case formula.

Actually measuring the Moon's parallax is more complicated than just described. The observer will not usually be on the equator and the Moon is stationary. These problems, however, can be overcome. Even the ancient Greeks were able to measure the Moon's parallax fairly accurately and hence know the Moon's distance in units of the Earth's radius (see Table 4.1 in § 4.7). The Moon's mean parallax from a modern determination is $57'2.60''$ which is nearly a degree. For the Moon one cannot use the small argument formulae and obtain a highly accurate distance.

The Sun's parallax is $8.794148'' = 4.263523\times10^{-5}$. This is quite minute and the small argument formulae would give a very accurate distance using it. The ancient astronomers were unable to determine the Sun's parallax; it was just to minute for their observational techniques (see Table 4.1 in § 4.7). In fact the accurate determination of the solar parallax is one of the long sagas of astronomical history.

It should be mentioned first that the solar parallax has never??? been accurately determined in the straightforward way we described above. Instead the distance to the Sun has been determined in some other manner and then the solar parallax obtained by inverting the sine case of equation (2.19). Thus the actual solar parallax values have no independent worth: they are just another way of writing the solar distance. Using a clever geometrical technique Aristarchos of Samos (c. 4th–3rd centuries BC) found the solar distance to be 360 Earth radii (Pa118–119, 497; Pe47–49) and hence a solar parallax of $9'30''$ (see Table 4.1 in § 4.7): this wildy too large due to Aristarchos' very poor observational data. Ptolemy gave a smaller, but still much too large, value of $2'51''$ for the solar parallax in the 2nd century AD (To265) and this value was still accepted by Tycho Brahe in the 16th century (Pa283). Kepler using Tycho's observations showed that the solar parallax must be less than $1'$ (Pa283). About 1630, G. Vendelinus used the method of Aristarchos of Samos with better observational data and obtained a value for the parallax of $15''$ (Pa283–284, 521).

2.6??? Cycles

One of the things most evident about astronomical periods is that they are usually practically incommensurate: i.e, any two of them will not usually form an exact ratio of small whole numbers. Physically, there is no reason why the should be commensurate in most cases. Since we can only measure periods to finite accuracy, large whole number ratios can be formed. However, these are not exact in principle since we cannot measure the periods exactly. Both large whole number ratios and cruder small whole number ratios were historically important in synchronizing astronomical clocks and in preparing ephemerides (tables of astronomical predictions) using cycles. In this context, a cycle is a specification of two sequences of two different astronomical periods. The number of periods in each sequence is a whole number: the whole numbers being the numbers from an approximate ratio. The total times of the two sequences are approximately equal of course. For example, given periods P_m and P_n where

$$\frac{P_m}{P_n} \approx \frac{m}{n} , \tag{2.21}$$

where m and n are whole numbers, the cycle period is approximately either of nP_m and mP_n. The cycle is more accurate for synchronization and prediction, the smaller is the relative difference

$$\left| \frac{nP_m - mP_n}{mP_n} \right| . \tag{2.22}$$

In this § 2.4.1, we will show how the year and day clocks are synchronized without resorting to a simple cycle. Then in § 2.4.2, we will show how cycles are used synchronize the year and lunar (i.e., mean lunar month) clocks and in § 2.4.3, the year and planetary synodic period clocks.

2.????.1 The Gregorian Calendar

The three basic astronomical clocks (which, of course, have incommensurate periods) are the day clock (measured the Earth's daily rotation on its axis), the lunar clock (measureed by the Moon's orbiting of the Earth), and the year clock (measured by the Earth's revolution about the Sun in a tropical year). The mean solar day is now defined to be 1 day in length or 86400 s. The lunar month is the synodic period of the Moon: e.g., the time from new moon to new moon. The tropical year is the time from spring equinox to spring equinox. The mean lunar month and mean tropical year are 29.530588 days and 365.24219878 days, respectively (Mo731).

Nowadays for civil purposes we are only concerned to keep the day and year clock synchronized. This is done by the Gregorian calendar instituted in 1582 under the patronage of Pope Gregory XIII (Pa220; No238). In the Gregorian calendar, years have either 365 or 366 days exactly: years of 366 days being called leap years. In its original version, years evenly divisible by 4, but not evenly divisible by 400 have 366 and all other years have 365 days. Thus 1995 has 365 days, 1996 and 2000 have 366 days, and 1900 has 365 days. The original mean Gregorian year (with symbol OGyr) is therefore exactly

$$\frac{400 \times 365.25 - 3}{400} = 365.2425 \text{ days.} \tag{2.23}$$

The original Gregorian calendar was actually not so good. In 3319.8 years, the original Gregorian calendar would be one day behind solar calendar. A modern revision of the Gregorian calendar stipulates that years whole number divisible by 4000 will also not be leap years (Ab129–130). Thus, the modern mean Gregorian year (with symbol Gyr) is exactly

$$\frac{4000 \times 365.25 - 10 \times 3 - 1}{4000} = 365.24225 \text{ days.} \tag{2.24}$$

With the revision, the modern Gregorian calendar will take 19,520 years to be one day behind the solar calendar. Thus, the Gregorian calendar has effectively solved the problem of synchronizing day and year clocks for the well beyond the foreseeable future of humankind. (The ideal calendar would be an even simpler and better solution, but it has been bypassed by history. See the note to Table B2 in Appendix B for the ideal calendar.)

2.????.2 Lunar Cycles

The problem of synchronizing the year and lunar clock has been abandoned by the international modern society for civil purposes. Certain cultures though remain interested in the problem. A major example being Islam which adheres to a year of 12 lunar months for religious purposes and needs to synchronize its calendar with the civil calendar (Gi46–47). In ancient times, however, many societies believed strongly in keeping the calendars synchronized: i.e., in maintaining a luni-solar calendar. Their reasons were mainly cultural and traditional. Once accurate solar reckoning is established, there is little practical benefit to most people in keeping track of the Moon.

In earliest time the synchronization of the luni-solar calendar was done haphazardly. Most years had 12 lunar months. If the seasons started to slip (e.g., the September harvest started falling in October), then an extra month, now called an intercalary month, was inserted. Religious or civil authorities determined when intercalation should be. As society became more organized, however, the need for a predictable intercalation was felt. Various lunar cycles of the mean lunar month and year came in to use to satisfy this need or perhaps just because they became known. We will demonstrate here how to create these cycles.

For the purposes of cycle making, we will use the Julian year of 365.25 days. This is a fairly close approximation to the tropical year and has the advantage of being exact by definition and allowing easy calculation of future times. Moreover, a synchronization of a lunar and Julian calendars is ideal if one is using a Julian calendar for dating. We effectively do use a Julian calendar for time intervals of 100 or 200 years (see § 2.???.1).

To modern accuracy the ratio, R, of the Julian year to the mean lunar month is 12.368531. Thus on average there are R lunar months per year. One can see that intercalations will have to take place every third year usually and sometimes every second year sometimes to keep the clocks synchronized. To determine an n-year cycle (where n is a whole number of years), one multiplies n by R to determine the mean number of lunar months in n years and rounds the number off to a whole number m which is the number of lunar months that most closely equals n years. Thus, one has a cycle of n years and

$$m = \text{Int}(nR + 0.5) \quad \text{lunar months,} \tag{2.25}$$

where 'Int' is the trunction-to-integer function. The accuracy of the cycle is determine by the relative error

$$E = \frac{\text{Int}(nR + 0.5) - nR}{nR} \ . \tag{2.26}$$

If E is negative, the lunar clock is running faster than the Julian clock; if E is positive, then the lunar clock is running slower. The smaller $|E|$ the better the cycle. The practical measure of error for any cycle is how many days off is it after one cycle period which for definiteness we take to be exactly n Julian years. Let this quantity be $E_{\rm d}$. The expression for $E_{\rm d}$ is

$$E_{\rm d} = Y_{\rm J} n E \ , \tag{2.27}$$

where $Y_{\rm J}$ is just the number of days in the Julian year. If the epoch when the epoch when the cycle was started was a new moon, then after exactly n Julian years the final new Moon of the cycle occurred a time E_d away in days.

As an illustration Table 1.1 shows lunar cycles and their diagnostics calculated using a modern value for the mean lunar month (29.530588 days) for some of the n values from 1 to 483. As n was increased from 1, only the cycles that improved in accuracy on (i.e., had smaller E than) all the proceeding ones were retained. If the mean lunar month were truly constant, we could find longer cycles that were more accurate still. However, there is in fact a secular increase in the lunar month of $\sim 4 \times 10^{-9}$ days per year (Appendix A, § A1). Unless accounted for this secular increase will cause a diminishing of returns in accuracy as one increases the cycle periods. I have stopped the calculation of cycles when the error caused by ignoring the secular increase became significant compared to the error in the cycle.

Table 2.1. Lunar cycles for the Julian year

n	m	d	E	$E_{\rm d}$	T_1
1	12	365.25	-0.02979588	-10.882944	-0.092
2	25	730.50	0.01062930	7.764700	0.258
3	37	1095.75	-0.00284576	-3.118244	-0.962
5	62	1826.25	0.00254426	4.646456	1.076
8	99	2922.00	0.00052300	1.528212	5.235
11	136	4017.75	-0.00039575	-1.590032	-6.918
19	235	6939.75	-0.00000891	-0.061820	-307.344
255	3154	93138.75	0.00000778	0.724552	351.942
274	3389	100078.50	0.00000662	0.662732	413.440
293	3624	107018.25	0.00000562	0.600912	487.592
312	3859	113958.00	0.00000473	0.539092	578.751
331	4094	120897.75	0.00000395	0.477272	693.525
350	4329	127837.50	0.00000325	0.415452	842.456
369	4564	134777.25	0.00000262	0.353632	1043.458
388	4799	141717.00	0.00000206	0.291812	1329.623
407	5034	148656.75	0.00000155	0.229992	1769.627
426	5269	155596.50	0.00000108	0.168172	2533.121
445	5504	162536.25	0.00000065	0.106352	4184.218
464	5739	169476.00	0.00000026	0.044532	10419.474
483	5974	176415.75	-0.00000010	-0.017288	-27938.454

Note.—The number of months and days in a cycle are m and d respectively. Note the day count d is exactly n Julian years, but the month count is rounded off to the nearest whole month. The E is the relative error in the cycle:

$$E = \frac{{\rm Int}(nR + 0.5) - nR}{nR} \ ,$$

where R is the number of mean lunar months per Julian year. The E_d is the number of days the cycle is off after one cycle period which we take to be exactly n Julian years. If E_d is negative, then lunar phenomena come early by $|E_d|$ days $|E_d|$ days. and if E_d is positive, then lunar phenomena come early by The number of Julian years until the cycle is desynchronized by one day is T_1 (assuming no secular variation in the lunar period). The cycles were calculated for n from 1 to 1000, but as n was increased only those cycles that improved on the proceeding ones were retained.

From Table 2.1, it is clear that shortest cycle with good accuracy is the 19-year cycle. This famous cycle is called (by moderns [Ne7]) the Metonic cycle after the Athenian astronomer Meton, who in 432 BC together with Euctemon discovered it (To12; Pa108; No65; Ne7). The Greeks did not put the Metonic cycle into immediate use, but did honor Meton with a statue (Ne7). The first known use of the Metonic cycle was by the Mesopotamians from 381 BC on (Pa52; see also Ne140; No65). It is not known if the Mesopotamians discovered the cycle before or after after Meton (Ne140).

The historical Metonic cycle is not quite what we have shown in Table 2.1. The historical Metonic cycle is 235 mean lunar months are equivalent to to 6940 days, not 6939.75 days (which is exactly 19 Julian years). This historical Metonic cycle is less accurate than the Julian-year Metonic cycle, since with it the lunar clock runs 0.31 days fast after one cycle period of 6940 days. Our Julian-year Metonic cycle is equivalent to the Kallipic cycle introduced by Kallipos of Cyzicus (c. 370–300 BC: Pe321), an associate of Aristotle's (To12), in 330 BC (To12; No65). This cycle equates $4 \times 235 = 940$ mean lunar months to $4 \times 6939.75 = 27759$ days which is exactly 76 Jyr (No65).

Besides their use for intercalation, lunar cycles allow a simple way of creating lunar ephemerides. To compute the predicted phases of the Moon for a given civil year, one merely writes down the observed phases of the Moon for a time one cycle period ago. As is evident from Table 2.1, this method of ephemerides calculation will not be too accurate for any cycle shorter than the Metonic cycle. To gain an order of magnitude decrease in relative error from the Metonic cycle (the Julian-year Metonic cycle) one needs to go to a 445-year lunar cycle. Planetary cycles (see § 2.???.3) allow planetary ephemerides to be calculated analogously.

Ephemerides calculated using lunar or planetary cycles (that use years) not only predict the synodic behavior (i.e., the behavior relative to the Sun), but also the sidereal behavior (i.e., the behavior relative to the fixed stars). This is for the following reason: in an integral number of sidereal years, the Earth returns to the same location relative to the fixed stars. Thus, after one cycle period of an integral number of Julian years (to continue using Julian years for clarity), the Earth has approximately returned to the same postion relative to the fixed stars and the celestial body has returned to approximately the same position relative to the Sun. Therefore the celestial body has also returned to approximately the same position relative to the fixed stars.

The above reasoning has a mathematical proof. Assume an ideal cycle of period P where

$$P = n_{\oplus} P_{\oplus} = n_{\rm syn} P_{\rm syn} \ . \tag{2.28}$$

Here $P_{\oplus}$ is the Earth's year, and since this is an ideal case we take the tropical and sidereal year to be exactly the same. The $P_{\rm syn}$ is the synodic period of a celestial body. For a celestial body (planet or Moon????), the synodic period is given by

$$P_{\rm syn} = \left| \frac{P_{\rm sid} P_{\oplus}}{P_{\rm sid} - P_{\oplus}} \right| \ , \tag{2.29}$$

where $P_{\rm sid}$ is the sidereal period of the celestial body. Substituting equation (2.29) into equation (28) and rearranging, we find

$$P = n_{\oplus} P_{\oplus} = n_{\rm syn} P_{\rm syn} = (n_{\oplus} \mp n_{\rm syn}) \, P_{\rm sid} \ , \tag{2.30}$$

where the upper and lower cases are for $P_{\rm sid} > P_{\oplus}$ and $P_{\rm sid} < P_{\oplus}$, respectively. It is easy to show that $n_{\oplus} \mp n_{\rm syn} > 0$ always. Thus P is also a cycle of an integral number of sidereal periods of the celestial body. Therefore, the sidereal behavior of the celestial body repeats after P as claimed.

A special use of the Metonic cycle in the Christian tradition is in the computation of the date for Easter (Pa217–221; No65–66; Ne7). It was early on established that Easter should be on the first

Sunday after the first full Moon of spring: i.e., the first full Moon after the spring equinox, usually March 21 in Gregorian Calendar. However, the greatest festival of Christendom could not be left to short-notice observational determination. A means of calculating the Easter date years in advance was desired. But the Christian Fathers were not too adept at astronomy, so some simple Easter reckoning procedure was required. With some complications, the procedure (called the *computus*) was essentially to use the Julian Metonic cycle along with a ad hoc correction called the *saltus luna* (Pa218). This Easter calculation procedure was due to Dionysius Exiguus in about 520, who also set the zero of the Christian (now also civil) calendar at the supposed year of Jesus' birth (see also Appendix C note). The difference between the Julian year and the tropical year, and the approximate nature of Metonic cycle itself lead to problems in Easter reckoning. With the Gregorian calendrical reform of 1582 (Pa220) these problems were cured, but addition of extra complications in the formula for Easter.

In creating Table 2.1, we used the Julian year as the unit for the year clock. It is of some interest to repeat the calculation using the tropical year since, like the mean lunar month, it is a true astronomical clock. Table 2.2 shows the lunar cycles for the tropical year: the table is exactly like Table 2.1, except for change in year unit and other defining quantities.

Table 2.2. Lunar cycles for the tropical year

n	m	d	E	$E_{\rm d}$	T_1
1	12	365.24	-0.02977515	-10.875143	-0.092
2	25	730.48	0.01065088	7.780302	0.257
3	37	1095.73	-0.00282446	-3.094840	-0.969
5	62	1826.21	0.00256567	4.685462	1.067
8	99	2921.94	0.00054437	1.590622	5.029
11	136	4017.66	-0.00037440	-1.504219	-7.313
19	235	6939.60	0.00001245	0.086403	219.899
182	2251	66474.08	-0.00001093	-0.726590	-250.485
201	2486	73413.68	-0.00000872	-0.640187	-313.971
220	2721	80353.28	-0.00000689	-0.553784	-397.267
239	2956	87292.89	-0.00000535	-0.467380	-511.361
258	3191	94232.49	-0.00000404	-0.380977	-677.206
277	3426	101172.09	-0.00000291	-0.294574	-940.341
296	3661	108111.69	-0.00000193	-0.208171	-1421.909
315	3896	115051.29	-0.00000106	-0.121768	-2586.893
334	4131	121990.89	-0.00000029	-0.035365	-9444.494

Note.—The symbols have the same meaning as in Table 2.1, except that n is the tropical year count, the day count is no longer exact, E is now calculated using $R = 12.368267$ (i.e., the mean number of lunar months in a tropical year), $E_{\rm d} = Y_{\rm T} n E$ where $Y_{\rm T}$ is the length of the tropical year in days (i.e., 365.24219878: see App. B, Table B2), and T_1 is now a count in tropical years.

We can see that up to a year count of 19, the same cycles with the same number of years appear in both Tables 2.1 and 2.2. Thus there is also a Metonic cycle for the tropical year. The tropical year Metonic cycle is also very accurate, although somewhat less so than the Julian year Metonic cycle. Above 19, the entries in the two tables are not the same. This longer, more accurate cycles are more sensitive to the difference in the lengths of the tropical and Julian years.

There is one other year, the Egyptian year, for which it is interesting to create a table of lunar cycles. The Egyptian year is exactly 365 days long. The calendar based on this year was used by the ancient Egyptians and later by astronomers down to Copernicus (Ne81; Ro130). A calendar with a year of invariant length and a whole number of days was found to be extremely useful in establishing an absolute astronomical chronology. To quote Copernicus:

> In computing the heavenly motions, however, I shall use Egyptian years everywhere. Among the civil [years], they alone are found to be uniform. For the measuring unit had to agree

> with what was measured. Harmony to this extent does not occur in the years of the Romans, Greeks, and Persians. With them an intercalation is made, not in any one way, but as each of the nations preferred. The Egyptian year, however, presents no ambiguity with its definite number of 365 days (Ro130).

Table 2.3 shows the lunar cycles for the Egyptian year: the table is exactly like Tables 2.1 and 2.2, except for change in year unit and other defining quantities.

Table 2.3. Lunar cycles for the Egyptian year

n	m	d	E	E_d	T_1
1	12	365.00	-0.02913135	-10.632944	-0.094
2	25	730.00	0.01132151	8.264700	0.242
3	37	1095.00	-0.00216278	-2.368244	-1.267
8	99	2920.00	0.00120829	3.528212	2.267
11	136	4015.00	0.00028891	1.159968	9.483
14	173	5110.00	-0.00023645	-1.208276	-11.587
25	309	9125.00	-0.00000529	-0.048308	-517.513
75	927	27375.00	-0.00000529	-0.144924	-517.513
311	3844	113515.00	0.00000511	0.580272	535.956
336	4153	122640.00	0.00000434	0.531964	631.622
361	4462	131765.00	0.00000367	0.483656	746.398
386	4771	140890.00	0.00000309	0.435348	886.647
411	5080	150015.00	0.00000258	0.387040	1061.906
436	5389	159140.00	0.00000213	0.338732	1287.153
461	5698	168265.00	0.00000173	0.290424	1587.334
486	6007	177390.00	0.00000136	0.242116	2007.302
511	6316	186515.00	0.00000104	0.193808	2636.630
536	6625	195640.00	0.00000074	0.145500	3683.849
561	6934	204765.00	0.00000047	0.097192	5772.080
586	7243	213890.00	0.00000023	0.048884	11987.562
611	7552	223015.00	0.00000000	0.000576	1060763.902

Note.—The symbols have the same meaning as in Table 2.1, except that n is the Egyptian year count, the day count is now exactly integral, E is now calculated using $R = 12.360065$ (i.e., the mean number of lunar months in an Egyptian year), and $E_\mathrm{d} = Y_\mathrm{T} n E$ where Y_E is the length of the Egyptian year in days (i.e., 365: see App. B, Table B2), and T_1 is now a count in Egyptian years.

The shortest high accuracy lunar cycle for the Egyptian year is the 25-year cycle. It is in fact somewhat more accurate than the Julian year Metonic cycle. The 25-year cycle was used by Hellenistic astronomers since they usually used the Egyptian calendar (Ne95). Even now the 25-year cycle has the convience of simplicity in casual (if not too important) calculations. For example, 1996 February 18 is a new Moon. Allowing for the 6 leap years, 2021 February 12 is predicted to be a new Moon. Because of the variation in the lunar month, this prediction may be off by a day or two?????.

2.????.3 Planetary Cycles

Planetary cycles for the year and planetary synodic period clocks are useful for writing planetary ephemerides as mentioned (see 2.???.2). This method for writing ephemerides was used by the ancient Mesopotamians as earlier as 523 BC (Pa55). In Table 2.4, we present for interest the simplest planetary cycles using Julian years again.

2.7??? Astrology

Astrology is not a monolithic thing. There is the broad distinction between Western and Eastern astrologies. The former arising in ancient Mesopotamia, and spreading and branching and cross fertilizing throughout India, the Islamic world, and the West. The latter initially entirely independent, but showing obvious signs of convergent evolution (No133). Other independent or semi-independent astrologies exist as well. But to the narrow the scope to the most familiar, we will just refer to Western astrology as astrology hereafter.

But even so confined, astrology is a broad domain. Modern astrology is the pseudo-scientific prediction of human affairs from astronomical events using a mishmash of ancient rules and ancient astronomical observations. It is practiced essentially for fun and profit, but very few take it seriously. One's daily horoscope is just a story for the day: an unreality check. Some take it much more seriously of course. This kind of astrology is not new—its at least as old as Babylon—but it is only a remnant of once mighty, if always problematic, lore.

In this section, we will just sketch a history of (Western) astrology. A full treatment is beyond our scope. We will touch on it again and again, however.

The disdain with which modern astrology is viewed by most people and assuredly by all scientists should not blind us to the fact that before about 1700 in Europe, astrology was serious philosophical discipline (No260ff, No270).

Astrology was

However, there are plethora of astrological traditions: they evolve, mutate, interbreed, often just stay the same, and often get made up brand new.

The origin of western astrology is in ancient Mesopotamian divination. The Mesopotamians (like many others) believed that the gods communicated information through portents, omens, and signs (which are all synonyms). A great many varieties of divination existed both unprovoked (e.g., a dog crosses your path) or provoked (e.g., you examine the viscera of dead animals to answer a question [extispicy]). The Mesopotamians systematized the techniques of divination in treatises and made large collections of omens. These works were considered the Mesopotamians and their neighbors one of the great Mesopotamian intellectual achievements (Op206).

Astronomical and meterological events (which were not always distinguishable) could also be omens. As early as the old Babylonian kingdom??? (c. 1900–1600 BC) (Op224, ????), the omen version of astrology was in evidence. Given the special majesty of the heavens it is not surprising that astronomical omens gained particular royal attention. The Assyrian kings of late Assyrian period (c. 900–600 BC????) had an astrological bureau??? to keep them informed (Op225, 227). These kings were terrible militarists and given the usual chanciness of war seemed particularly in need of inside information???. From those times on there were often to be astrologers about the courts of kings. This tradition in Europe continued up to the 17th century. Perhaps the tradition has not ended yet: there is a persistant story that Nancy Reagan consulted astrologers (No267).

That astronomy (i.e., systematic and systematizing study of astronomical events) gained great impetus of astrological needs seems clear. One also needs to remember the human tendency for studies to become ends in themselves: so may it have been for astronomy even among the Mesopotamian astrologers. By the 8th century BC, a much higher level of astronomy was being practiced. The high point of Mesopotamian astronomy was achieved after 300 BC (see § ???).

It is possible that horoscopic astrology was initiated in Mesopotamia in the 5th and 4th centuries BC (Op225), but the point is disputed. Horoscopic astrology, as it has developed, is the prediction of human affairs (and also human medical conditions) based on the astronomical state at the time of birth (or even conception). This birth state stamps the character and, in medical astrology, the body of a person. In extreme form, the person whole life is determined, but few astrologers have ever held that doctrine. The horoscope, itself, is both a chart showing the position of the celestial bodies in some or other arcane coordinate system at a particular time and the interpretation of this chart. The natal horoscope for a person is the key one. Of course, later astronomical events also influence a person, but these are also interpreted in terms of the birth astronomical state.

Actually, almost all astronomical states are completely determined: something the Mesopotamian astronomer-astrologers learned quite well. Therefore, almost all the astrological influences a person will receive are foreordained. However, working out all influences in advance was tedious and for the astrologer's clients expensive. Often a horoscope would be bought for only so many years in advance. Comets, of course, were unpredictable and so added some uncertainty to the equation.

In the Hellenistic period Mesopotamian astronomy and astrology entered the Greco-Roman world. Whether or not horoscopic astrology originated in Mesopotamia or not, it certainly flourished in the Greco-Roman world. Unlike the older astrology, horoscopic astrology was democratic: it applied to all, high and low. It became a common coin of the realm. However, then as now there were skeptics. One ancient (Cicero?????) remarked that astrologers must smirk at each other when they meet.

There was a plethora of astrological and astronomical literature of various sorts most of which has not survived. Some of it was evidently very crude and continued long after better astronomy and more sophisticated astrology developed (No119).

At the highest level was the *Tetrabiblos* of Ptolemy (c. 100-175 AD: To1), one of the greatest astronomers. This book is the masterpiece of astrology, and so recognized by the the most enlightened of subsequent astrologers. Here astrology is presented as a science shorn of its anthropomorphic??? elements. It has to be emphasized, however, that Ptolemy distinguished clearly between astronomy and astrology: the former was an exact science; the later was approximate, speculative, but sublime.

Ptolemy was a rare spirit; most professional astrologers, however, would just crank out horoscopes and interpret them according to the rules. They never needed to look at the stars, but merely consult tables of dubious accuracy. Nothing much has changed since. The astrological rules gives the influences of the various astronomical events, but weighting of the influences is an art. Obviously, a good personal astrologer can use their intimate knowledge of their clients to tailor their horoscopes to the person. The rules just applied blindly give no statistically valid correlation between horoscope and person: a resulted repeated endlessly.

Astrology was driven out of the Christianized Roman empire of the 4th and 5th centuries (No122–124), but revived a bit in the 8th century (No122–124). Also in the 8th century astrology flowered in the Medieval Islamic world. Islamic astrology then flowed into Europe in the 11th and 12th centuries. Astrology became deeply deeply embedded into European culture. Although there were people who disapproved of it from a Christian perspective, it was generally tolerated.

The practitioners and their conceptions of astrology were various. Medical and meteorological astrologies were sciences to be used and developed (No262, 269). Personal astrology

In the Middle Ages it returned to the Western world and became quite respectable despite its basic incompatibility with Christianity. Most Medievel and Renaissance astronomers were astrologers. Of course, there were many astrologers who could not be called astronomers. Physicians, for example, used astrology for medical reasons.????

The last great astronomer to be an astrologer was Johannes Kepler (1571–1630). For most of his career his job was Imperial mathematician (i.e., astrologer) in the service to the Holy Roman Emperors. Kepler knew the ordinary rules of astrology were worthless, and told his patrons as much even when supplying them with the predictions based on those rules. But he held fast to the belief that there was a connection between microcosm and macrocosm: a hallowed correspondence between the God-given soul and the divine handiwork above. He cast many horoscopes for himself seeking his own significant explanation. He hoped for a purified astrology. He was a transitional figure. A generation after Kepler, astronomers and astrologers had gone their separate ways. Nowadays astronomers do not cast horoscopes (and mostly do not know how) and astrologers take no interest in astronomy: it has been a divorce made in heaven.

To finish this section, a brief word on the zodiac. The zodiac can mean three things. First, a belt about 18° wide centered on the ecliptic. The orbits of all the planets, except Pluto, are confined to this belt (Cl69). The second meaning is the twelve traditional constellations that straddle the ecliptic. (The constellation Ophiuchus also crosses the ecliptic but it is omitted from the zodiac [Ze10].) The twelve zodiacal constellations are listed in Table 1.1. Most of these constellations are animal figures: only Libra is clearly inanimate. Hence the word zodiac which is derived from the Greek zoidiakos, circle of animals (Ba1420). The zodiac constellations were mostly derived from the ancient Mesopotamians who first came up with the idea of the zodiac. From them the zodiac passed to the Greeks and from the Greeks to us.

The third meaning of the zodiac is the twelve signs. These are the signs that turn up in astrology. The signs are in fact nothing but 30° intervals of the ecliptic that are measured starting from the spring equinox. Two thousand years ago, the spring equinox was in the constellation Aries: that was the Age of Aries. The twelve signs were named for the the zodiacal constellations they contained at that time. Due to the precession of the equinoxes, the spring equinox shifts 1.396° west per century. Thus it has shifted 27.93° in 2000 years. Currently, the spring equinox is in the constellation Pisces, but sometime soon (or late) it will enter Aquarius. Since the borders of constellations are variously defined the timing of the dawning of the Age of Aquarius is moot. Due to the precession a sign now contains, more or less, the constellation just to the west of the one that gave the sign its name.

A person's astrological sign is the sign that the Sun is located in at the time of their birth. Thus a person born in the period March 21–April 19 is an Aries. For most of the March 21–April 19 period, the Sun, however, is in the constellation Pisces (i.e., Pisces is in solar conjunction). The linkages of the signs with the solar calendar which is almost exactly the civil calendar makes things easy for modern astrologers. The most basic element of a person's horoscope can be read off a calendar without any reference to an astronomical table. Table 1.1 gives the natal period for each sign.

Table 2.4. The zodiacal constellations and signs

Constellation and sign	English name	Approximate mean date of solar conjunction of constellation	Natal period for sign
Aries	Ram	April 30	March 21–April 19
Taurus	Bull	May 20	April 20–May 20
Gemini	Twins	July 5	May 21–June 20
Cancer	Crab	July 30	June 21–July 22
Leo	Lion	September 1	July 23–August 22
Virgo	Virgin	October 11	August 23–September 22
Libra	Balance	November 9	September 23–October 22
Scorpio	Scorpion	December 3	October 23–November 21
Sagittarius	Archer	January 7	November 22–December 21
Capricorn	Fish-Goat	Februay 8	January 20–Februay 18
Aquarius	Water Bearer	March 27	Februay 19–March 20

Note.—The approximate mean dates of the solar conjunctions of the constellations are from Ze10. Since any source is as good as any other for the natal periods for the signs, the ones given are from "Breszny's Real Astrology" in the *Nashville Scene* (Breszny 1996, p. 58–59). Like the planets, the zodiacal constellations and signs have traditional symbols (Ne228). However, they cannot be reproduced by TEX.

Needless to say, for most people astrology is just playing at believing. Your daily horoscope gives you a story for the day.

3. Ancient Mesopotamian Astronomy

Ancient Mesopotamia is where the first significant advances over prehistoric astronomy were achieved. It is in Mesopotamia that we have the first detailed records of astronomical observations. (I do not count observational records implied prehistoric structures since that would be stretching the meaning of the term record and because these cannot be dated precisely and are not all that detailed in fact.) Moreover, the first mathematically predictive theories of astronomy developed in Mesopotamia relying on the Mesopotamians great achievements in mathematics.

Before delving into the history of astronomy in Mesopotamia, we give an introduction to Mesopotamian history (through to the year 100 AD). This is profitable just for general interest (since popular knowledge of ancient Mesopotamia tends to be slight) and for background information to the astronomical achievement. Indeed, the background information is very relevant. One cannot have records for astronomy or history before there was writing, and writing most probably began in Mesopotamia (Po52, 56). The Mesopotamian mathematics provided the tools for predictive astronomy. And the Mesopotamian propensity to divination provided an incentive for astronomy in the service of astrology.

3.1 The Land

Mesopotamia is a Greek name meaning land between the rivers??? (Ba763, 948). The term was used by Greek and Roman writers (Llo12) and is not a translation of a name used by the ancient Mesopotamians themselves. It is, however, usefully descriptive. The centers of ancient Mesopotamian civilization lie between or close to the rivers Euphrates and Tigris that have their sources in the Taurus and Zagros mountains (more or less in modern Turkey and Iran respectively (Llo12). Euphrates and Tigris are names that appear in the earliest texts and thus date back to a literal time immemorial (Po173). The Euphrates starts in Turkey, flows through western Syria, and then south through Iraq to the Persian Gulf. The Tigris, to the east of the Euphrates, does the same, except that it barely touches Syria. About 150 km from the Persian Gulf, the rivers merge, or more accurately lose their identity in a coastal land of marshes and lakes (Llo14–15).

The north and east of Mesopotamia are enclosed by the aforementioned Taurus??? and Zagros mountains (Llo12) These mountains were outside of ancient Mesopotamia (as we think of it), but not disconnected from it at all. Trade and military incursions from and to the mountains occurred???. To the north-west the Euphrates extends to within 200 km of the Mediterranean (Po12), giving the Mesopotamians a connection to that realm of history. To the south and south-west are the Syrian and Arabian Deserts (Po12; Llo13). Approaching Mesopotamia from the west, one encounters an escarpment with a fall of $\sim$ 30 m to the plain of the river valley; this escarpment provides the name Iraq, meaning 'the cliff" in Arabic (Llo12–13). The deserts were inhabited by nomads as far back history extends (Po4). However, the deep desert was probably nearly impassible before the domestication of the camel which does not seem to have been before circa 1000 BC (Po5).

In Mesopotamia four regions might be generally distinguished based on geography and climate. First, are plains fringe the north and north-east by the foothills of the Taurus and Zagros mountains (Po11). This fringe has sufficient rainfall for agriculture: i.e., more than $\sim$ 300 mm per year (Po13). The exact contour line of sufficient rainfall has varied over history and so has the border of settlements. It is in this region that Assyria (the major northern component of Mesopotamia) was located.

The second region, south of the northern fringe, east of the Euphrates is the Jazirah ('the Island'), a an arid limestone plateau, the home of nomads in ancient and modern times (Llo14; Po4–5). The ruggedness of this area and its elevation make it unsuitable for for irrigation.

The third region, south-east of the Jazirah is the alluvial plain of the Euphrates and Tigris. It starts about 600 km from the coast (Llo14–15). However, even 500 km from the coast the land is less than 20 m above sea level (Po6). The rivers have, of course, created this plane by deposition over a geological time span. Because of low contrast in the plane the rivers have changed their courses several times in past few thousand years (Po15). These are easily undestood. The alluvium from river of shallow grade tends to be deposited along its course. Gradually banks (levees) built up about the sides of the river especially from flood deposit. In time the river runs higher than the surrounding plane Then a large flood can cause it to burst a bank, and take a new course.

The alluvial plane is arid and the unwatered parts are sand or barren desert (Po15). The summer daytime temperatures are in the range 43–54 degrees Celsius range and there is no rain for 8 months (Llo17). There are rainstorms in the winter, but they are insufficient to lift the aridity. In a natural state only along the river edges would there be extensive vegetation despite the natural fertility of the alluvial soil. However, from prehistoric times irrigation has been practiced.

The original farmers expanded along the river edges, but found they could easily water the surrounding land by building canals off the rivers (Po174). Because the rivers actually flow above the surroundings at least in flood time, all that was needed was to cut through the banks to get a flow of water. The problem was to control the water. Too much water would drown the land. Moreover, the rivers are in flood in sometime April through June, but the growing season is over the winter (Llo17). Therefore a complex network of canals and reservoirs was needed to control and store the river water. Naturally too, the canals and reservoirs would tend to silt up and need period redigging or relocation (Llo17). Another problem is salination of the soil. Even fresh water has a salt concentration and when continually evaporated from a field will leave salt deposit that ruins the fertility of the soil. Moreover, the ground water also rises through because of irrigation and it is salty likewise (Llo18). Proper drainage and other techniques can limit the salination problem and the ancients eventually implement some of these methods (Llo18). Still salination was a recurring problem and there are today areas white with salt deposition that cannot be farmed.

The activities of centuries of irrigation works have created a complex network of dry canals (Llo17). These can be seen from the air and give the impression of a land once having been densely and extensively cultivated. This is somewhat illusionary because only a fraction of these areas were in use at any one time (Llo17). Silting up of canals, salination, and the changing course of the rivers caused human settlement to shift. Human disasters could also change the land usage. The constant warfare of the late 9th and 10th centuries AD (the late Caliphal period of early Islam) with its constant extortions from the farmers ruined much of the region by causing depopulation and making investment in agriculture projects worthless. The economic activity retreated at least to a degree back to nomadic pastorialism (La133–136). However, despite the damages to the alluvial plain caused by nature and human activity, it seems to have remained remarkably productive throughout ancient times and only gradually suffered lowering productivity over the millennia (Llo17).

The fourth region of Mesopotamia that we distinguish here is the coastal marsh land. The ancient Akkadian name (see § 3.2 for Akkadian) for this region literally translated is the Sealand (Op404). Today as in ancient times, people live in the marsh land and have adapted to it (Po6–7; Kr87ff). The modern marsh dwellers live in reed houses on islands that can often be flooded and travel about on boats which traditionally at least are just canoes. They rely on fishing, rice cultivation, and the water buffalo husbandry. Their ancient counterparts must have led a similar life, but rice cultivation was introduced only about 1000 BC (Kr92) and the history of the water buffalo in the region is complicated. The modern water buffalo were introduced in the early Islamic period. (Po7). However, an ancient introduction of water buffalo to Mesopotamia (although not certainly to the marsh land) in the time of the Akkadian dynasty (see § 3.2) seems to have occured, probably from the Indus civilization (Po164–165). This introduction known from art works seems to failed since the animals disappear from art after the Akkadian dynasty.

The modern marsh dwellers probably in fact have no historical connection to the ancient marsh dwellers (Po7). The marsh has probably been repopulated since the beginning of the Islamic period. But because of convergent social and economic evolution, the modern and ancient marsh dweller probably live very similar life styles.

3.2 Outline of Political and Cultural History

3.2.1 Early Development

The prehistory of the Mesopotamia and surrounding areas has been intensely studied. Here we only say that in this region the Neolithic phase of humankind began with the introduction of agriculture and animal husbandry. The transformation to the Neolithic was probably evolutionary and dispersed (Llo26) and probably began before 8000 BC (Llo28, 30).

The Neolithic transformation came comparatively late to alluvial plain of southern Mesopotamia; perhaps circa 5000 BC (Po23). However, it is in this region that the first urban culture developed probably in the 4th millennium BC (Po73). The course of this development is still far from fully understood, but it seems to have been a local evolution. There is strong evidence for a cultural continuity from the period of 5000 BC onward.

The city of Eridu (now Tell Abu Shahrain), now $\sim$ 250 km from the sea and south of the Euphrates was ruputed by the inhabitants in the earliest historical time, the Sumerians, to be the oldest city in the world (Llo15, 39). Archaeologists do not deny the possibility. At the lowest level of excavation on virgin sand, there is a small temple dated to 4900 BC recognizably the forerunner of great Mesopotamian temples of later times (Llo39, 41–42). The temple is about 3 meters, and its plan shows features that can be identified as a cult niche at the back (the seat of the god) and a central offering table (Llo41–42; Po118–119). This basic layout would persist for the inner sanctums of Mesopotamian religion. Above this simple structure, in upper layer of excavation, are a series of temples culminating in what was probably a massive structure on a raised mound of earth (Llo39; Po25).

Eridu's prehistoric role as a shrine was probably one of the determining causes for a large settlement to develop there; a settlement with economic and political importance, and complex specialization of labor (Po73–74). Archaeology and the historical record agree that the major settlements that were spread over Southern Mesopotamia had strong patriotism based on their city, its temple, and its protecting deity (Po26). Thus almost other cities probably evolved similarly about shrines ancient shrines even if they do not go or cannot be traced back as far as at Eridu. The these other earliest of cities include Ur (the birth-place of the Patriarch Abraham in the Bible [Kr35]), Uruk (the city of the legendary [Kr114ff], but also historical Gilgamesh [Llo92]), Sippar, Kish, Lagash, Adab, and others.

The people who built these cities were the Sumerians and perhaps the Akkadians????, who appear to have been present in the early 3rd millennium BC (Po36). The Sumerian language has no known relatives and the origins of the Sumerians has been debated. Some, such as Kramer (Kr33), have contended that they arrived in Mesopotamia only in the 4th millennium BC from perhaps central Asia based on philological interpretation of place names (Po24; Llo62–63)). However, the continuity of archaeological evidence particularly in regard to religious practices back suggests strongly that the Sumerian culture developed locally (Llo62–64). A compromise possibility might be that the Sumerians were invaders, who imposed their language, but who adopted the local culture (Po24). But in fact invaders who impose their language are in the minority in these

3.2.2 Writing

3.2.4 The Hammurapi Dynasty of Babylon

****************** needs heavy revsion.

Hammurapi (Hammurabi) (r. 1792–1750 BC [Op337]) was the 6th king of the 1st dynasty of Babylon (Op337; Po39) which we will call (as many do) the Hammurapi dynasty (e.g., Ne29) Prior to the Hammurapi dynasty, Babylon had been a small town whose existence is known of only as far back as the Ur III dynasty (i.e., probably since circa 2100 BC) (Op155, 336). The Hammurapi dynasty lasted 1894–1595 BC (Op337; Po39). It was founded by Amorite invaders from Syria, who also established other regimes in the northern Fertile Crescent (Ll157). The Amorites had been nomads (Po42–43); their conquest is part of a recurring pattern in history of invasion of settled areas by fierce rim-land dwellers, who subsequently acculturate or even assimilate with their subjects. The Amorites spoke a Semitic language (Op49; Po36), but the Amorites of the Hammurapi dynasty seem to have largely adopted Old Babylonian, the language (also Semitic) of their subjects (Op54–55).

In the early years, the Hammurapi dynasty did not rule extensive territory and may not always have been independent (Op156). Under the father of Hammurapi, Sin-muballit, and Hammurapi, Babylon grew into an empire ruling Mesopotamia from the Persian Gulf to circa 1000 km inland (Op156; Po44). Hammurapi's most famous exploit (in modern eyes at least) was his sake of the city and palace of Mari (capital of a rival Amorite kingdom) in circa 1757 (Llo157–159; Po49). The sack left Mari (modern Tell Hariri [Llo157]) an archaeological treasure-trove: a palace preserved as well as could be expected and an archive of cuneiform records (Po49, 141).

Hammurapi is also famous, of course, for his code of laws preserved on 42 stelae or columns (Po 288). The exact significance of the Code of Hammurapi is much debated: a true set of laws that the king expected to be obeyed or a literary idealization of the way things ought to be done (Po289; Op158) Postgate contends that the Code is something of a mixture: some provisions are repeats of old laws and others are royal reforms (Po289). As king ruling over a recently enlarged kingdom, it is reasonable and just that Hammurapi should try to establish a new standard code of laws to hold throughout his domain.

Subsequent to Hammurapi, the Hammurapi dynasty seems to have entered a slow political decline: the 1st ynasty of Sealand (a dynasty from the marsh lands on the Persian Gulf it seems [Op404]) took over the old southern Sumerian cities and Kassites from the north-east harassed the Babylonians (Llo160). In 1595, King Mursilis I of the Hittites swept south into Mesopotamia and sacked and burnt Babylon, putting an end to the first dynasty (Llo160). The Hittites, however, quickly withdrew, and a Kassite dynasty established itself in Babylon. The Kassites assimilated rather completely to the Babylonian culture, and thus ensured its continuance (Llo160; Op158).

The Hammurapi dynasty overall established the fame of Babylon that has never vanished, even its power ended with the Persian conquest of 539 BC (Llo222-223) and it is now an archaeolocial site only. The period of the Hammurapi dynasty is culturally important since it preserved and augmented the literary tradition of the old Akkadian and Sumerian periods (Llo157). It is in fact the best documented period of Babylonian history with many records opening windows on political, social, and economic life (Op154).

It may also have been the high point of literacy in the Near East before the introduction of the alphabet (Po69). The Hammurapi dynasty is also the period from which we have our first window on the Mesopotamian achievement in higher mathematics; we do not if this mathematic developed then or earlier, only that it existed then (Ne29–30).

We should note that the dates of the Hammurapi dynasty and all of exact Mesopotamian chronology before about 1500 BC are uncertain (Op335; Pa34) These dates have been established by a synchronization with the Venus cycle as observed in the time of King Ammizaduga, 2nd to last of the Hammurapi dynasty (see § 3.5). But the astronomical calculation is uncertain because of the repeating nature of astronomical phenomena by an additive multiple of 64 or 56 years (Pa34). Thus historical research must establish the absolute zero. The chronology given here has been long favored (Pa34–35; Op337; Po39), but an alternative is also possible. In the alternative chronology, 64 years is added to all dates (No29): e.g., Hammurapi's reign becomes 1728–1685 BC.

3.3 Cosmogony and Cosmology

3.4 Mathematics

3.5 Omens and the Birth of Astrology

3.6 The Astronomical Achievement of the Late Period

In the late period of Mesopotamia (c. 300 BC–75 AD: the Seleucid-Parthian period) we find the appearance of a "consistent theory of lunar and planetary motions" (Ne97): i.e., a mathematical and predictive theory which for brevity we will call the late theory. But we do not know exactly when or how or by whom the late theory was developed. The creative phase has been almost deleted from history. What we have is the evidence of about 300 tablets from the late period: less than 250 ephemerides and about 70 procedure texts (Ne105–106). These tablets show the theory as already complete (Ne102).

Some plausible surmises can be made about the time frame in which the late theory developed from the timing of known astronomical innovations. Before 530 BC (just after the start of the Persian period) there was no regularity in the pattern of years with an intercalation of an extra lunar month. From 530 BC on, however, the 8-year lunar cycle (8 years and 99 mean lunar months) (see Chapt. 2, Table 2.1) was in use (Pa51–52). This cycle is not very good: mean lunar predictions are off by a day after 5 years from a synchronization, and this is less than the cycle time itself (see § 3.10???, Table 3.1). After 380 BC, the Metonic (19-year) cycle (19 years and 235 mean lunar months) (see § 3.10???, Table 3.1) was in use (Pa52). This may represent a considerable step in exact astronomical knowledge, since the Metonic cycle is a factor of about 50 more accurate than the 8-year cycle and, in its Julian year version (which of course the ancient Mesopotamians probably did not use), is only a day off in mean lunar predictions after 307 years.

?????????????????????????????????????? Still to be revised below.

The latest astronomical text is dated to 75 AD which is nearly at the end of cuneiform writing altogether (circa 100 AD: Op352). By the late period, the Mesopotamian civilization had long lost political independence (due to the Persian conquest in 539 BC ??? and later other conquests) and in the telescoping retrospective view of history can be seen to be near to its end. That the highest level astronomy was reached in this period demonstrates, if any demonstration is necessary, that various cultural achievements are not always closely connected.

One of the early signs of improved astronomy is the use of the 19-year Metonic cycle of lunar intercalation (see App. A, Quiz 2.4 and App. B, Table B3) after 380 BC as determined from the intercalary years of the Persian and Macedonian kings (Pa52). A more powerful indication is the existence of ephemerides based on long periodicities.

Synodic phenomena for a planet (e.g., conjunctions, oppositions, and retrograde motions) depend on the phase of the synodic cycle (the events that occur in the course of synodic period). However, the synodic cycle is not in fact a exactly repeating cycle. Even if a planet moved in a circular orbit in the ecliptic plane with uniform speed, the position on the Celestial sphere (i.e., relation of the fixed stars to the planet) will be different at the same phase in succeeding synodic periods. And of course those conditions do not obtain. The exact synodic behavior (length of the synodic period, and celestial position, speed, and ecliptic longitude as functions of synodic phase) depends also on the sidereal cycle of the planet. Thus, the planet's overall behavior will repeat over a period that is a common multiple of the synodic and periods (hereafter the repeat period): i.e., when the planet is located at the position relative to the Sun and fixed stars. Thus we want integer solutions for the factors m and n in the equation

$$mP_{\rm syn} = nP_{\rm sid} \ , \tag{3.1}$$

where $P_{\rm syn}$ and $P_{\rm sid}$ are the synodic and sidereal periods, respectively.

There can be no exact common multiple. The (mean) synodic and sidereal periods are known only to finite accuracy and if we use all the digits of modern accuracy available, we would find impractically long repeat cycles in any case. Thus, no ideal pair of integers can m and n be found. However, one can use approximate, practically useful repeat periods. Then the ephemerides for a planet for a given year can be predicted approximately by simply reporting the ephemerides one repeat period back.

The Mesopotamian astronomers used this method together with clever corrections to account for the lack of perfect repeat periods (Pa54). Of course, they also had to take account of the purely chronological difficulties caused by the variable years of luni-solar calendars. As an example, the Mesopotamians found a repeat period for Jupiter of 81 (tropical) years (Pa54). This period constitutes $76P_{\rm syn}$ (modern $P_{\rm syn} = 398.9$ days: F565) and and $7P_{\rm sid}$ (modern $P_{\rm sid} = 4333$ days: F565). To (text book) modern accuracy, the ideal ratio of $n/m = 3989/43330 \approx 0.09206$: this ratio is the number of sidereal periods per synodic period. The Mesopotamian ratio of

$n/m = 7/76 \approx 0.09210$ which is not too bad. Obviously, the longer the history (time baseline) of accurate observations, the more accurately the repeat periods could be determined and the more accurate the resulting ephemerides.

Another periodicity that the Mesopotamians probably used (as early circa 380 BC) was the Saros period (Pa58–59, but see the caution of Ne142). The Saros period is the nearly exact repeat cycle of eclipse phenomena. The word Saros is from the Babylonian word *shar* meaning universe (or something similar) or the large number 3600 (Ne141). Its association with the eclipse cycle is a philological??? error due to Edmund Halley (of the famous comet) in 1691 (Ne142). The length of the eclipse cycle is 223 mean lunar months or ≈ 6585.32 days (18 tropical years and 11 days, not counting fractional days). The remainder of about 1/3 days means that at each new Saros period the Earth has rotated about 120°. Three Saros periods constitutes a nearly exact replication of all eclipse phenomena. However, because the Saros period is not an exact periodicity (Pa60), solar eclipse tracks can effectively wander all over the Earth given enough time (Gi??? and references therein that badly need to be consulted). The evidence for the Mesopotamian use of the Saros period is the 'Saros Canon', a cuneiform tablet investigated by Strassmaier and Epping (Pa60).

One has to emphasize, that knowledge of the Saros period and the methods described below were insufficient to predict solar eclipses (Ne119). Lunar eclipses can be seen by all people located on the night side of the Earth. Thus, lunar eclipse prediction based on simple means is feasible and the Mesopotamians achieved this. Solar eclipses can only be seen in a geographically restricted track; totality in a only very narrow track. The Mesopotamians never developed the geometrical or geographical insights to predict the tracks of solar eclipses. All they could do was to exclude an eclipse or predict its possibility (Ne119).

The Mesopotamians also had a more sophisticated approach to prediction than the periodicity method. This was an approximate difference method (Ne110–113) used from 250 BC (Ne115). All of the ancients (including the Greeks and everyone before circa 1500 AD?????) failed to understand the concept of velocity in a straightforward way: i.e., as the ratio of a distance to a time. However, the Mesopotamians did calculate the angular distance travelled by celestial bodies in time intervals (e.g., a lunar month) as a function of position on the Celestial sphere and time. Given the fixed time interval, these angular distances (distance differences) are effectively velocities. The Mesopotamians, however, further simplified their differences by effectively making a linear approximation between their periodically recurring minimum and maximum values. Consequently, the differences as function of time form a zigzag function when plotted rather than a smooth sinusoidal curve (Pa66). There is no evidence that the Mesopotamians ever used graphical representation however?????.

Using tables of differences (and also similar, but simpler approaches Ne114) quite accurate ephemerides could be calculated in principle. However, because of errors in the differences, errors in the table predictions could be a few degrees???. The practical purposes the periodicity method could provide somewhat more accurate ephemerides, but still with maximum errors of a few degrees. The Greek astronomers did no better. Only with the *Rudolphine Tables* (1627) of Kepler based on Tycho Brahe's observations and his own planetary model, were maximum errors usually reduced to a degree (e.g., Ze55). One should note that accuracy of only 5 % means that celestial bodies could be a fist (at arm's length) off predictions. Probably one could still find the celestial body with this accuracy provided it was not too faint, but true times of conjunctions, oppositions, occultations, etc., could be missed by days???? without keeping watch near and in advance of the predicted times.

It is possible that the Mesopotamian had what, anachronistically, we could call a rather positivistic theory of celestial phenomena aside from theological beliefs. Thus what one sees in the sky (motions in two dimensions on the Celestial Sphere with their fixed cycles) was the theory itself. This can certainly be called a valid scientific theory. If you did not see what you saw, the theory would be falsified. The parameters of the theory needed fitting and the fitting can be improved by repeated observation. There is a cycle of observation from the real world, adjustment of the parameters to yield better predictions, renewed observation, and so on. The cycle is effectively infinite: ever improving predictions heading to a limit of exact prediction. This is essentially the scientific method applied, except for the search for generalization. They did not try generalize so far as we know. They could, for example, have tried to bring terrestrial notions of geometry and motion into their theory: if successful this would have obtained some unification of earthly and celestial physics and so have created a theory that we would say was more fundamental.

However, if the Mesopotamians had some non-theological picture beyond just what was seen, then it has not been recorded and it never became the basis for empirical study: i.e., it was never submitted to the scientific method. Perhaps sometimes they did image the planets in three dimensional space, but nothing came of it. The theological picture that the gods moved the celestial bodies and fixed their periodicities was certainly part of their view, but since that picture is not falsifiable (however true or false it may be), it is not a scientific theory. They did come to believe that the interference of the Moon caused solar eclipses, but never that that of Earth caused lunar eclipses?????. The Earth was a nearly flat, immobile surface at the bottom of sky, not an object on the sky dome. This cosmological picture could certainly be treated as a scientific theory has discussed in Ch. 2.???, but it never was.

And who were the late Mesopotamian astronomers? The cuneiforms tablets that we have come from two archives found in Babylon and Uruk (Ne136): those from Babylon rarely have colophons. Colophons are the final sections of the tablets containing usually the title (often an incipit) if there is one, the name of the scribe, the owner of the tablet, the date, and sometimes a serial number (Op241; N17). The colophon might say "Tablet of A" (Ne136) and give a genealogy of A. There can be an invocation for divine aid or a curse on the illicit borrower. The Uruk texts indicate that the writers came from two scribal families both claiming a priest as an ancestor. But we do not know if the scribes were astronomers or mere copyists???. From Greco-Roman writers (Pliny, Strabo, and Vettius Valens) the westernized names of three Mesopotamian astronomers are known. One of these, Kidenas may be the Kidinnu mentioned in some of the colophons.

Although we do not know the exact channels, Mesopotamian astronomical conventions and data did pass into the Greco-Roman world. This is not surprising since the Mesopotamia was part of an Hellenistic kingdom from Alexander's conquest (330 BC??) till the Parthian conquest (181 BC????). Even before, Greek travellers (Herodotos???), traders and craftsman (????), and mercenaries (Xenophon???) had penetrated the Mesopotamian world. Hipparchos (c. 190–120 BC) used some Mesopotamian data (Pe49; No93ff). From shortly before Hipparchos, in the *Anaphorics* of Hypsicles, we find the first Greek appearance of the Mesopotamian division of the circle into 360° and sexagesimal arithematic (No93). From him and other sources these data and conventions have been passed onto posterity along with some constellations, the Zodiac???, and the concept and raw materials of astrology.

3.7 Postlude: The 24-Hour Day and the Egyptian Year

The achievements of the Egyptians in mathematics and astronomy were minor compared to those of the Mesopotamians (Ne72, Ne80, Ne91). This does not, however, imply a general backwardness in sciences. In medicine, both modern commentators (Op296) and Herodotos (see Op299) put Egypt in ahead of Mesopotamia. For astronomy, however, there are two areas in which Egyptian influence is worth some discussion here: the 24-hour day and the Egyptian year.

Our 24-hour day is an Egyptian relic. Its evolution in Egypt is somewhat complex. By the time of Seti I (c. 1300 BC) the day was divided into 10 hours (which reflects a basically decimal number system; N85–86) to which two hours were added for the twilight periods of morning and evening, respectively (Ne86). The night was also divided into 12 hours from as early as 2100 BC (Ne88) about the time of transition between the Egyptian Old and Middle Kingdoms (Op348). This division also had its root in the decimal number system. The path from a decimal number system to 12 is rather intricate; Neugebauer (Ne81–86) has given a description of this path, but note that a crucial argument seems to be omitted (Ne85). Thus, the 12 hours for night and day sum to give a 24-hour day. Originally, the lengths of the hours was uneven, but Hellenistic times at least seasonal hours had come into use (Ne86). Seasonal hours divide both day and night into 12 equal parts???; the length of the day thus depends on the time of year. Only on equinoxes are all the hours of day and night of equal length. These equal-length hours or equinoctial hours were originally only used by Hellenistic astronomers (Ne81). The 24-day with seasonal hours passed into the Greco-Roman culture in Hellenistic times and was from there passed onto Islamic and European cultures.

Probably with the introduction of mechanical clocks in the 13th century in Europe (Co???, Pe???), the equinoctial hours replaced seasonal hours in general use????. The minutes and seconds of our time reckoning came from the Mesopotamian number system and were also introduced in Hellenistic times (Ne81). Thus our time reckoning is a hybrid system.

From early historical times (c. 2800 BC) the Egyptians used a solar-based civil calendar (Pa82–83; Mo82-83). Their year had exactly 365 days and was divided into 12 months of 30 days each with the left over 5 days (epagomenes or epagomenal days) being placed at the end of year. The epagomenal days were festival days and were considered unlucky for work. The Egyptians also used various lunar calendars for the regulation of some festivals.

The civil calendar was not synchronized with the tropical year: there were no leap years and no intercalations. Because the Egyptian year is shorter by about a quarter day than the tropical year, the Egyptian calendar would run ahead of a solar calendar and would eventually cycle through the whole tropical year. The divergence of the two calendars was sufficiently slow, however, as not to pose any major problems. A person who lived to be over a hundred would only find a difference of about 25 days between the seasons and dates of their infancy and old age: e.g., the May blossoms of youth would bloom in the June of final years. Since the tropical year is 365.24220 days long, a complete cycle would take $365.24220/0.24220 \approx 1508.0$ tropical years. Pannekoek (Pa83) gives $365/0.25 = 1460$ tropical years which cannot be the cycle period. What he is really trying to arrive at is the Sothis period, the time interval between the heliacal risings of Sirius (the Egyptian Sothis) on the same solar date. For the Sothis period he gives 1456 tropical years taking into account the proper motion of Sirius. The only conclusion I reach is that the 1460 value is just an illustrative Sothis period assuming the true tropical year is 365.25 days, there is no precession of the equinoxes, and Sirius has no proper motion. Thus the true Sothis period of 1456 accounts for these three factors and is indeed the time interval between the heliacal risings of Sirius on the same solar date. However, the Sothis period is not the cycle period, except that Pannekoek also says that it is. So there is some error someplace.

Although the Egyptians did not need to intercalate lunar months, they did have a cycle analogous to the 19-year Metonic cycle for predicting lunar phenomena. There cycle was a 25-year cycle (see App. B, Table B3 also) where the years are Egyptian years, not tropical years as with the Metonic cycle. The number of mean lunar months in 25 Egyptian years is

$$\frac{25 \times 365}{29.53059} \approx 309.0016 \tag{3.2}$$

or almost exactly 309. The time it would take for the mean lunar phase to run one day ahead of a 25-year cycle prediction is approximately $1/0.0016 \approx 62$ cycles or ~ 15000 years. Lunar phase predictions from the 25-year cycle will, of course, have errors of up to ± 2 days??? because of variations in the lunar cycle.

4. Greek Astronomy

4.1 The Setting

The world of the ancient Greeks is the littoral and the body of the Mediterranean Sea (midland or inland sea) with a few, but important, journeys beyond. The Greek homeland, Greece proper (Hellas to the Greeks who to themselves were Hellenes), is a complicated peninsula jutting out into the north side of the eastern Mediterranean and the copious surrounding small islands. There are limited fertile plains separated by rugged uplands. The land of the Mediterranean is often quite dry and to those from wetter climates looks somewhat barren. The second and third life bloods of the Mediterranean world are olive oil and wine: this true today and in Antiquity.

As has long been noted, the separation of Greece into fertile islands and real islands has promoted disunity and also a florescence of individual community cultures. However, using the highway of the Mediterranean communication and trade among Greek cities and other cultures was easy. Even bulk goods could be traded relatively cheaply by sea which was not the case on land for among other reasons, the horse-strangling nature of the ancient horse-harness (Gies28, 32, 46–47). Mediterranean sailing is much less demanding than oceanic sailing. Usually ships could set courses within sight of land??? which was a great advantage when only primitive non-mathematical navigation was possible. Tides are measurable, but practically unnoticeable. Wharfs could be set almost at water level and probably taverns too just as in our times.

The collapse of the Bronze-Age Mycenaean Greek culture before 1000 BC caused a migration of some Greeks to the coast of western Turkey which became their Ionia. The Mycenaean collapse ended literacy and left the Greeks in what we call their Dark Age. The end of the Dark Age is loosely sometime around 800 BC by which time there are independent Greek cities all over Ionia. In addition, from during the Dark Age and for 2 or 3 centuries thereafter???, both the Ionian and old Greek regions were both active in founding colonies in Italy, Sicily, Southern France, Eastern Spain, and around the Black Sea. The Italian and Sicilian colonies were particularly extensive and active politically and culturally until the Roman conquest circa 300–200 BC left them a backwater. Colonization, naturally, implied conquest of territories from the indigenes. The colonization period is called??? the Archaic Age; it is generally taken to last down to the time of Marathon in 490 BC. It is in that age, in the Mediterranean setting that the Greek achievement in science begins in Ionia.

4.2 The Origin of the Presocratics

By definition, the Presocratics are Greek philosophers who flourished before and, illogically, while and even somewhat after Socrates. They were deeply interested in natural philosophy, unlike Socrates who relegated natural philosophy to lesser place relative to ethics at least in regard to his own investigations. For our purposes, natural philosophy can be defined as the study of the natural world based on reason and observation with a goal of trying explain phenomena in terms of elementary principles or laws. To elaborate, the reliance on elementary principles means that phenomena were assumed to occur in cause and effect chains. The causes could be mechanistic meaning that their effects are incidental or teleological meaning that the cause is ordained to bring about the effect. Note that mechanistic in this context has no connection to machinery and that teleological does not imply a cause has in any sense an intelligence or an awareness of the effect. The Presocratics tended to a mechanistic view of causation. Aristotle is the prime example of teleologist, but, of course, he is not a Presocratic.

The Presocratics can be discussed only with caution. None of their own writings survive in completeness; only quotations and commentaries from latter writers have come down to us. Nevertheless, plausible reconstruction of some of their views are possible (e.g., Fu). They were not very empirical. In general they performed no experiments and made no quantitative or detailed observations. An experiment reported by Empedocles (Pe134) is exceptional. For the most part, they attempted to understand the universe using only casual observations and reason. Such an approach may indeed be possible; indeed it may be the a way to reach absolute truth about nature. It is, however, obviously very prone to error and has been far less successful than the scientific method at yielding adequate theories. However, the achievements of the Presocratics are not negligible. They are the starting point for both western philosophy and science.

The broad philosophical aspects of the work of the Presocratics is outside our scope. However, their cosmological speculations are a part of astronomical history and relevant for later astronomy and science in general. A restricted, retrospective view of the Presocratics, is that they were theory creators for a cycle of the scientific method that stretched over millenia. We will discuss some of the Presocratics in § 4.3. Here we will try to understand why they rejected anthropomorphic mythical explanations for natural phenomena and pursued natural philosophy instead. Our argument is given under six headings.

1) Like most cultures, the Greeks emerge into history with a set of myths accounting for natural phenomena. And like many cultures their myths were anthropomorphic. The sky god and king of the gods was Zeus, the thunderer and also father of gods and men. The sea had a controlling god, Poseidon. The Sun and Moon had gods (Apollo and Artemis or, alternatively, Helios and Selene). And there were a host of other gods, controlling human destinies, inflicting punishments, ripening the corn. Hesiodos and Homer and a wealth of community tradition provided the post-Dark Age Greeks a sufficient mythological account. Why should that or any anthropomorphic mythological be unsatisfying?

Well, skeptical rejection of anthropomorphic divine action is hardly ruled out by the world as humans perceive it. The mythical beings who are supposed to control the world are capriciously human. But the world seems too regular in some respects for the vagaries of merely human nature. In other respects it seems too chaotic for humanlike beneficence or malice. Ordinary human experience shows that mindless chains of cause and effect must drive most natural happenings: i.e., things behave according to natural causes. Day to day economy, for example, would be impossible if it were not so. Why not make the mental leap to saying natural causes are all causes, except those that spring from obviously animate beings. Once the mental leap is made to this natural philosophy, then it along with some cleverness can be used to explain at least some natural phenomena. That not all natural phenomena can be explained can plausibly taken to be a lack of insight on the part of the explainer rather than a failure of natural philosophy.

That the above argument was apparent to the Presocratics seems almost to be necessarily so. They probably were not alone. Even from the Mesopotamian civilization there are some accounts of skepticism about some aspects of anthropomorphic divine action (Op226–227). These accounts end by the unbelievers being undeceived of their skepticism, but those endings are morals. It is plausible that a practical if not a theoretical skepticism was not rare in Mesopotamia or in any society with anthropomorphic mythological religion. It is also unexceptional to say that skepticism in a general sense is a common human attribute. We can conclude by saying skepticism of anthropomorphic mythological explanations is a necessary, but not sufficient condition for natural philosophy.

Another, condition at least helpful to natural philosophy was the lack of enforcement of doctrinal orthodoxy among the Greeks. Honoring the religious customs of the city or other pertinent unit was more important than adhering to mythology????. Only one Presocratic, Anaxagoras, was ever prosecuted for impiety (Fu173). And in that case there was also a political motive to the prosecution????. Instances of persecution of impiety alone must have been at least rare.

We should add here that the Presocratics were not at all necessarily atheists. Some may have been so. Certainly, Plato thought some at least had atheistic opinions (Fu173ff). However, at least in some cases, the example par excellence being the Pythagoreans, they were undoubtably seeking a different and in their minds truer religion than that of tradition (see § 4.4). Another example is Xenophanes of Colophon who spoke of "One god, greatest in the company of gods and men, like mortal men neither in his shape nor in his thought," who "Without toil by the thinking of his mind he shakes the universe (quotes from Fu162)."

2) There are factors that could have an enhanced the skepticism and freedom from traditional

religion in the Presocratics: their the economic and social conditions. The first Presocratics were urban Ionians. At that time, Ionia was the most active maritime colonizing and trading areas of Greece. Ionians would have had contacts with other Greeks and cultures all over the Mediterranean, including the Egyptian, Phoenician, and Mesopotamian cultures. Such contacts can only have made the parochialness of mythical accounts rather apparent. Additionally, the breadth of the Ionians travels would have shown the errors of any purely local conceptions of the world. The Oceanic voyages of the Europeans from the 15th century on had the same effect on an even grander scale.

The cosmopolitanism acquired by some Greeks in their socio-economic life is a plausible aid to skepticism. However, the Phoenicians were in a 'similar boat,' and did not grow skeptical and philosophical. More explanation is needed.

3) Conceptional boldness in the Greeks may have been supported by their sense of being a young and free society. That they must have had this sense at least by comparison to the eastern and Egyptian cultures is clear. They knew those cultures had much longer histories and that they had much to learn from them. But they were not self-abasing. Their immense sense of civic loyalty and liberty, and their heroic past as recounted by Homer, made them seem to themselves vigorous as compared to the conquest-weary, tyrant-laden centuries of the east and south. Thus, they could adopt what they like from the east and south and pride themselves on liberty of action and fearlessness. Such cross-cultural eclecticism could support an eclecticism of thought.

4) The pluralism of the Greek world must also be considered. There was not one capital, one cultural orthodoxy, one model. The Greeks were unconsciously in an experiment of numerous trials. With so many trials, the probability was increased that somewhere in the Greek world natural philosophy (and many other cultural features) would arise.

5) Perhaps the key basis for the development of natural philosophy was the liberty of speech and emphasis on reasoned argument that developed in the Greek cities states, the poleis???, as suggested by Furley (Fu167). Although, most states were not democratic, there was usually some participatory assemblies?? for a wide class of enfranchised citizens (i.e., those who were male, not enslaved, and usually of some economic standing). The smallness of Greek states (as opposed to the vast empires) and their lack in general of deadening unmovable political authority meant that what was said in assembly, in the law courts, in the agora (marketplace, but maybe too much of a synonym with assembly???) was not perceived as pointless and at least often not as dangerous. Thus, the arts of persuasion and reasoned debate were honored and held to be significant.

That meaningful dialectic could expand its empire from the political and social worlds to the natural one is eminently plausible. We now believe that free speech and debate are key ingredients in the development of modern culture (which is certainly technologically and scientifically the highest ever and arguably morally the highest at least in our conceptions). That it performed similarly for the Greeks may be unamazing.

6) To the cultural conditions, one must add the accidents of human genius and the tendency of investigation to have an autonomous existence once started. These extra ingredients are essential. Total reduction of intellectual achievement to cultural conditions (such as in orthodox Marxism) is now generally believed to be an invalid approach and at least as used by some (e.g., some Marxists) unscientifically unfalsifiable. In fact if one accepts an objective knowable reality and that the form of investigation leads to some approximation of truth, then a complete reduction of natural philosophy to cultural conditions is ruled out absolutely.

With the recipe of cultural conditions, accident, and autonomy of intellectual development, we may have a plausible explanation for the emergence and florescence of natural philosophy. Nevertheless, we are speculating: we have no clear record of intellectual development of Presocratics and none at all for the earliest. Moreover, even if we did, we could never know the causes with the surety of a well known physical result. As discussed in Chapt. 1, that is not within historical research's power.

4.3 The Presocratics (except for the Pythagoreans)

The Presocratics were not specialists; they were attempting establish grand systems embracing all natural phenomena and the entire universe. They were not very empirical and relied on casual observation and rational argument. They did not practice the complete scientific method and cannot be classified as scientists in the modern sense. The whole scope of their speculations and logical arguments are beyond our scope. We will only discuss a few Presocratics and their ideas without their much of their justifications. For us, their ideas are important as starting points, often unrealized or realized only much later, for empirical investigation. All of our discussion must be with the caution that the writings of the Presocratics have all been lost, except for fragments. Thus, their ideas can only be learned or deduced from what later writers said about them or from the fragments. Misinterpretations and gaps have to accepted.

4.3.1 Thales, Anaximander, and Empedocles

The first name is Thales of Miletos (c. 625–545 BC). Miletos was one of the most active Ionian maritime cities. He became a semi-legendary figure and only a little of what is attributed to him is reasonably certain. It seems certain, though, that he said that water was the primal element from which everything else developed and that the Earth was a disk floating on water (Pe12). However, the story that Thales predicted a solar eclipse is certainly untrue: there was no theory then or for centuries that would have permitted this (Ne142). Neugebauer (Ne148) thinks it very unlikely that Thales discovered the geometrical proofs attributed to him later. The importance of Thales for history (aside from contributions which may have been lost) is the introduction of the notion of elements.

Anaximander of Miletos (c. 610–547 BC), a pupil of Thales (Pe13), has a slightly more complete record. Attributed to him, among other ideas, are the following (taken from Pe14–15). The stars, Moon, and Sun are really wheels that filled with fire. What we see as the celestial bodies in the sky are apertures through which light and heat pass. Stoppages of the Moon and Sun apertures cause eclipses. The wheels are at different distances from the Earth: from highest to lowest (i.e., farthest to closest) the Sun, Moon, stars, and planets. Whatever the justification for the wheel idea, it is a beginning point for Greek thinking about the celestial bodies in three-dimensional space. The wheel idea may, in fact, have borrowed from Persian cosmology (Pe15). Anaximander's Earth is like a stone pillar and we reside on one of its end. He speculated that living things were generated in moisture when this was evaporated by the light of the Sun and some sort of evolution has occurred leading to humans.???

4.3.2 Parmenides of Elea and the Spherical Earth

Parmenides of Elea flourished in the 1st half of 5th century BC (Pe374). His home town Elea is an ancient city on the coast of Lucania in south-west Italy (Ba387). In about 450 BC he visited Athens (Pe374). He is the earliest Presocratic from whom sufficient fragments of writing survive that his philosophy and arguments can be understood with some confidence (Fu31). He wrote his results in a poem (poetry proceeds prose) and some of this poem is quoted in surviving works by Simplikios of Cilicia (6th century AD) (Fu31, 50; Pe391–392). Parmenides purely philosophical points are not our subject, but he is considered to be important and original in these (Fu31). Here we merely report his conclusions on cosmology which are likewise important and original.

In Mesopotamian and earlier Presocratic thinking the Universe has a fundamental linear direction along which up and down are measured (Kr99; Fu53). The surface of the Earth is flat and perpendicular to the fundamental direction; gravity points down the fundamental direction. The Mesopotamian cosmology has an over-arching sky dome (Kr99) and Anaximander has his rings, but the fundamental direction is primary. Parmenides is reported to have been the first to have declared the Earth to be spherical (Fu41, 56). This is a striking advance to say the least. How did he come this this conclusion?

Furley believes that the result was arrived at from metaphysical speculation on what the universe must be like (Fu56). A fragment of Parmenides' poem evidences this:

> But since there is an outermost limit, it is perfected
> from all sides, like the mass of a well-rounded ball,
> equally balanced from the centre everywhere. For neither greater

> nor smaller must it be in one place or another.
> For neither is there Not-Being, which might stop it reaching
> its like, nor is there Being such as to be
> more than Being here, less there, since all is inviolate.
> For equal to itself from all sides, it lies uniformly in its limits.
> (quoted from Fu54).

Furley believes that Parmenides is talking about the universe as a whole and stating that the universe is spherically symmetric and has a central focus (Fu54). This implies a spherically symmetric Earth if the Earth is taken to be at the center. Since things fall to Earth, why does the Earth not fall: because it is already in the center. Furley admits Parmenides own words are somewhat obscure (Fu54). But in Plato's *Phaedo*, Socrates reports that an unnamed person has convinced him of the spherically symmetric universe and a spherical Earth in the center and Socrates' words are reminiscent of Parmenides' (Fu55).

Furley contends that the evidence for spherical Earth and a centrifocal universe emerged later (Fu56). However, if Parmenides did come upon the spherical Earth from purely metaphysical speculation, then I contend it was a lucky strike. It seems to me more plausible that observations at least partially led Parmenides to a spherical Earth and universe, and the philosophical argument was then partially an a posteriori explanation. First, the heavens do look spherical and to first approximation celestial bodies revolve around the Earth. Parmenides, who had proposed an absolute monistic theory of reality (Fu38), might be led then by analogy to contend that the Earth too must be spherical to be consistent. It is also possible (although there is no evidence) that Parmenides was aware of some or all of the physical proofs of the spherical Earth first mentioned by Aristotle (Pe45):

1) As one moves north and south the altitudes of the celestial bodies and the stars change in a manner consistent with a spherical shape for the Earth.

2) The shadow of the Earth on the Moon during a lunar eclipse is round no matter how exactly the Moon, Earth, and Sun are aligned.

3) At great distances at sea masts can be seen when hulls are beneath the horizon. (This observation takes sharp eyes and excellent weather conditions).

The thing that makes it especially plausible for me that Parmenides could have known these proofs is that he evinced considerable insight into astronomical theory. It is reported that he was the first to claim that the Morning and Evening Star were the same object, Venus (Fu56). (Some, like the Mesopotamians, may have guessed this earlier, but in there planetary calculations [which come from after Parmenides] they treat the morning and evening appearances of Venus and Mercury as those of different objects [No54–55]. Odysseus in the *Odyssey* was unaware of the identicality of the Morning and Evening Star [Pe11]). From his own words we know that Parmenides knew that the Moon shone by light reflected from the Sun and he may have been the first person in history to record this idea. The 2nd of Aristotle's proof for a spherical Earth would not have been possible for someone who did not know this.

As a final word, the Pythagoreans also have a claim to being first to have proposed a spherical Earth. Like Parmenides, they give no empirical argument for their proposal (Pe18), but they may have had one. They seem to have argued for this shape for the Earth and the univserse on the grounds of its perfection (Pe51)

4.3.3 Empedocles and the Elements

The idea of elements that began with Thales led to the development of several systems of elements. The one that endured, with some modifications, was that of Sicilian Empedocles of Agrigentum (c. 484–424 BC [Pe331–332]), who, like Parmenides, wrote in verse: "my tale is not aimless or unknowing" (from Fu96). His was the system of the four elements: fire, air, water, and earth (Pe124). All other substances were supposed to be made of these in some combination. These elements were conserved: i.e., never destroyed or created (Pe124). Their combinations could change and thus account for observed material transformations. It seems that fire, in particular, was not always what we mean by fire, but sometimes also invisible light and heat?????. To the four basic elements he added love and strife which are 'forces', but are sometimes treated almost as if they were elements also (Fu84); love and strife did not become part of the standard dogma of the four elements.

The elements were considered to be continuous (i.e., infinitely divisible) and the notion of a vacuum was excluded by most Presocratics and later by Aristotle (Pe133) and those who followed him strictly. Aristotle began one of his anti-vacuum arguments with the punning remark that "the so-called vacuum will be found to be really vacuous" (from Fu190). From somewhat different sets of arguments, Plato and Aristotle would add a fifth element, the ether (meaning upper air or sky [Ba412]), out of which the celestial objects were made (Pe126, 128). In Aristotle's view, this element was unchanging and eternal, and unmixed with the four elements of sublunar world (Pe128). Until the 17th century????, the four elements would dominate the interpretation of chemical knowledge. From our perspective, it is easy to see that the four elements could not have led to much chemical enlightenment.

4.3.3 The Atomists

An interesting dissent from the view that there were four elements and no vacuum was the atomic theory (atomism) of Leucippos (possibly of Abdera, Elea or Miletos [Fu115]) (5th century BC [Pe365])) and Democritos of Abdera (c. 460–370 [Pe328]). *Atomos* is an adjective meaning uncut, indivisible, and unmown (as of grass) (Fu123). Because the lack of surviving whole works and the fact that Leucippos and Democritos were associates, their independent ideas cannot be discerned (Pe131). They postulated that all matter was made out of invisibly small atoms that were eternal and indivisible and of infinitely many kinds (Pe131; Fu123). Combinations of these atoms made the materials and bodies that we see. There ideas came from the observation that all material things are perishable, therefore the imperishable must be below the level of perception (Fu124). To be perishable or breakable there must be void in an material (Fu124). They gave the following illustration. When one cuts a apple, the blade goes between the atoms, it cannot divide them???.

Space (the universe or *to pan*, the all) was infinite and a void, except for the atoms moving and colliding continuously (Fu136, 140ff). The continuous motion is natural for the atoms in the void (Fu146). Spontaneous vortices could be set up out of which condensed infinitely many worlds (*kosmoi*) (Fu140). The vortices provided a mechanically way of sorting material into different concentric shells (Fu142). The still center of the vortices and revolving outer regions could provide a crude description of a geocentric planetary system. Very probably the vortex idea came to the atomists (and other Presocratics who also employed it: e.g., Empedocles Fu92) from simple observations of whirlwinds and whirlpools???. Because the vortex-spun worlds are only made of atoms, they are perishable too. Our world (the visible universe or cosmos) is only one of the worlds. The Earth was a disk and was probably seen as floating on air (Fu143). The Celestial sphere was a surrounding spherical membrane separating us from the chaotic regions beyond (Fu143). The reconciliation between the disk Earth with linear (i.e., uni-directional) gravity and the spherical shape of the cosmos (our *kosmos*) is not clear (Fu143). The vortex idea would seem to lead to only to an axial symmetric *kosmoi*. The notion of a membrane surrounding the cosmos may have been drawn from biology, in fact embryology (Fu143).

In later times, Epicuros of Athens (341–271 BC) adopted the atomism with some modifications. Two of these were that there were only a finite number of kinds of atoms though infinitely many individuals of each type and that the atoms could combine by means of small hooks (Pe132). Epicuros was one of the grand philosophic system builders. The atomism formed the natural philosophy part of his system. He and his school, the Epicureans (mistakenly associated with simple hedonists), were not empirical chemists. The atomism (with its spontaneous and mechanical elements) was philosophically acceptable to their relative atheism in which the gods are far off and do not interfere.

Given the closeness of the atomism to modern views (their atoms to our atoms; their vortices to our galactic and solar system formation), it is interesting to speculate on what would have happened if the atomism had become the primary ancient doctrine rather than the four element theory and the finite cosmos of Aristotle (Fu189; Pe89). Perhaps chemistry would have advanced much more rapidly given a fortuitously correct starting point. The vortex theory would have been had to have been modified, of course: a spherical Earth would have been an early necessity and the disappearance of the Celestial sphere membrane eventually. But the case is that atomism was a minority view and was strongly criticized by Aristotle, among others, who is also one of the principle sources for our knowledge of atomism (Fu136). Atomism was not forgotten, however: the idea passed through the Epicureans (especially the Roman Lucretius [c. 98–55 BC] and his famous philosophical poem *De Rerum Natura* [Pe133]) to 16th century scientists (Gassendi, Charleton, Boyle, and Newton) to Dalton with his empirical chemical atomism, the beginnings of modern atomism (Fu123). Thus, Greek atomism provided a starting point (certainly non-essential) for the modern atomic theory. Nevertheless, it is probably true to say that the ingeniousness of Greek atomism was not matched by its historical importance.

4.3.4 A Presocratic Summary

From the limited point of view of the scientific method, the Presocratics were mostly stuck in the theory creation stage. Much of their theorizing was not falsifiable given the practice and sometimes not even falsifiable in principle. However, the arsenal of ideas they provided is remarkable. Certainly the boldness of their thinking contrasts with the over-rigid adherence to authority that would characterize natural philosophy in later times from circa 200 BC down till about 1600. Harnessed to the full scientific method, the Presocratic ideas might have resulted in rapid scientific progress. However, in the ancient world the full power of the scientific method was never generally realized. It is also possible that a continuous scientific progress which we have become accustomed to since circa 1600 was not sustainable in the Greco-Roman civilization (see Sect. 4.12).

4.4 The Cosmology of Pythagoreans: The Philolaic System

The most celebrated of the Presocratics was Pythagoras of Samos (c. 550–500 BC) (Pe385). He was certainly a real person as near contemporaries refer to him (Fu49) He is supposed to have left his native, an island in the Aegean, and travelled in the Near East (Pe16ff). Later Pythagoras is supposed to have settled at Croton in southern Italy and founded a school or a movement. The movement certainly existed and even held political power at times in Croton and Metapontum (also in southern Italy) in the 5th and 4th centuries BC (Fu50). The movement was ascetic and believed in a physical and mental regime to overcome the physical body. The Pythagoreans practiced vegetarianism and believed in reincarnation. They also believed in enlightenment through mathematics (Pe16). Exactly, what mathematical discoveries they actually made is hard to be certain of. Their own traditions ascribed many to Pythagoras himself, but this may be a result of a tendency to expand the reputation of a founder and increase the venerability of a discovery. The prime example of a discovery claimed for the Pythagoreans is, of course, the Pythagorean theorem, but that was empirically well known 13 centuries earlier at least in the Old Babylonian kingdom (Ne36). They may have re-discovered it or, perhaps, provided a geometric proof. There is also the discovery that a vibrating string's frequency is inversely proportional to its length. (Note, I am using anachronistic and non-musical terms here. See Pe17 for a musical discussion of this discovery.) This discovery has never been forgotten: at least some of nature's laws are mathematical. The Pythagoreans were not, however, inspired by this discovery into an experimental search for mathematical relationships, but into an arbitrary occultism of numbers (Pe18).

In a startingly innovation from the Mesopotamians, the Pythagoreans and later Greeks saw a need to give depth to space and treat celestial motions as tracing out geometrical shapes in it. The Pythagoreans also made the Earth spherical possibly because of the mystical notion of the perfection of the spherical shape (Pe51), but perhaps they were already aware of some of the arguments for a spherical Earth adduced by Aristotle (Pe45 and see below). From Pythagorean circles comes the one of the earliest geometrical models of the cosmos with a spherical Earth; it is mostly known from accounts by Aristotle and Aëtios of Antiocheia (c. 100 AD) (Pe52; Fu57–58). It is probably later than Parmenides (Fu57), but is more detailed. The attached name is Philolaos of Croton or

Taretum (late 5th century BC).

In the Philolaic system there are 10 concentric celestial spheres, which are not the celestial bodies (i.e., planets, Sun, etc.), but shells on which these bodies reside. The idea of these celestial spheres may go back to Anaximenes of Miletos (c. 540 BC) (Pe51), another Ionian Presocratic. (We use small 'c' in celestial spheres to distinguish them from the Celestial sphere of the fixed stars which is one of their class, of course.)

At the center of the Philolaic system is neither Sun nor Earth, but the Central Fire: applied to the Central Fire were the epithets 'the Hearth of the Universe,' 'the House of Zeus,' 'the Mother of the Gods,' 'the Altar,' 'the Meeting House' and 'the Measure of Nature' (quotes from Fu57). Around the Central Fire are ten concentric celestial spheres carrying the celestial bodies in a uniform circular motion (Pe54). The order of the spheres going outward is the Anti-Earth, Earth, Moon, Venus, Mercury, Mars, Jupiter, Saturn and star (Celestial) spheres. The Earth revolves (on its sphere) daily around the Central Fire (counterclockwise from a north pole view or 'eastward'), but always always has the Greek hemisphere opposite the Central Fire (Pa100): hence this object is never seen. The Anti-Earth revolves also in a day and is on the opposite side of the Central Fire (Pe52; Fu57): it too is never seen. The introduction of the Central Fire and the Anti-Earth are unnecessary, except for mystical reasons. Aristotle says that the Anti-Earth was introduced so that the number of celestial spheres would be 10, a perfect number to the Pythagoreans (Fu58).

The daily eastward revolution of the Earth sphere accounts for the daily westward circling of the Celestial sphere. In a sense the Earth must be rotating on its axis relative to the fixed stars (nearly fixed stars: see below) in order to keep the Greek hemisphere pointed away from the Central Fire. However, one should probably think simply of the Earth as being rigidly attached to the Earth sphere. The outer celestial spheres carry the celestial bodies eastward on the ecliptic, and thus account for prograde (i.e., direct) motion of these bodies. The Sun orbits the Central Fire in a year naturally. The Celestial sphere was given a small imperceptable motion, not to account for precession which was unknown, but to raise the number of moving spheres to the mystical 10. There is no explanation for retrograde motions. Unless non-uniform motions are allowed, there seems no way to account for why Venus and Mercury are never far from the Sun. We know of no attempt to make the system mathematically predictive.

Eclipses were accounted for by Sun and Earth interposition just we would. In a lunar eclipse the Anti-Earth and Central Fire would also be on the Sun-Earth-Moon line, unless some small disalignment of these bodies were allowed. Such a disalignment would spoil the sysmetry of the system, however.

The Philolaic system was probably never intended as completely serious astronomy (Fu58), but only as a mystical system: perhaps the designation science fiction world is not totally anachronistic either. It is nevertheless an influential and interesting system. One main interesting feature about the Pythagorean model is that there was no compunction about having the Earth move and be off the center of the cosmos: there is no astronomical point to this motion and no consideration of the physical implication. The notion of a moving, non-central Earth would be become a very minority view in Greek astronomy as we will see below.

An influential element of the Philolaic system is the reinforcement of the idea (probably due to Parmenides: Fu53ff; see also 4.3) a universe with a central focus, even if that central focus was not the Earth in the Philolaic system. Centrifocal universes would become the accepted norm at least from Eudoxos on (see § 4.5).

Another influential element, possibly original to the system, is the introduction of the celestial spheres that carry the celestial bodies along. These celestial spheres, sometimes called crystalline spheres, would be retained by Aristotle (and be made of his ether: Pe128) and from him would be passed on as long as Aristotelian cosmology reigned. Eudoxos and Aristotle would make the spheres strictly geocentric (see §§ 4.5 and 4.6), but the epicycle system-builders would not retain that restriction. It seemed more realistic to the Greeks and all their followers till the 16th century to have the celestial bodies carried by physical objects however invisible. Tycho Brahe showed that there were no solid celestial spheres (still excepting in some minds the Celestial sphere itself) in the late 16th century because comets passed right through where they were supposed to be (Th257–258, 262). (He was inspired by the calculations of Michael Mästlin and the ideas of Christoph Rothmann, to be fair all around [Th257–258].) In some minds, Tycho's for instance (Th306), the Celestial sphere itself continued to have some solidity for awhile longer.

The celestial spheres can be considered as an example of physical analogy (or physical intuition) being led astray. The 19th century ether (not to be confused with its classical namesake) is another famous example (see § 8.3.1).

The Pythagorean principle of uniform circular motion would also be retained for along time as an absolute principle: honored by Plato (see § 4.3.5), honored in the breach by Ptolemy (see § 4.10), redeemed by Al-Shātir in the geocentric world (see § 4.4.3), redeemed by Copernicus for the heliocentric universe (Ro4, 10), and lastly retired by Kepler in his *Astronomia Nova* (Ca134), but still with honor (Ca270).

An influential and strange feature of the Philolaic system was the harmony of the celestial spheres: a music generated by the moving of the celestial spheres which we cannot hear being inured to it since birth (Pe53–54). The idea of the harmony of the spheres would resound down the centuries. Kepler in a much evolved, sophisticated, and mystical way would encorporate the harmony of the spheres in his theory of world harmony given in his book *Harmonice Mundi* (Ca264ff, esp. Ca269–270).

Lastly, a science fiction feature, reported by Aëtios, is that the Moon is supposed to be inhabited by animals fifteen times more powerful than terrestrial animals (Fu58).

4.5 Plato and Eudoxos

Plato's (427–347 BC [Pe378]) philosophy has an importance for science in general and astronomy in particular that must be kept distinct from its value in philosophy itself. Here we are only concerned with its scientific importance. A principle aspect of Platonism is doctrine of ideas. There is a real world, not the material world that we see, that consists of eternal ideas which are the real objects and one of kind and are the real objects of thought (Pe22ff). Material objects are perishable copies of the ideal objects. We have an inborn understanding of the ideal world and only this allows us to understand the material world by recollection (Pe24). Mathematical objects are somehow not the same as ideal objects being multiple, but eternal and perfect. However, deductive geometry and mathematics (still very new fields in his day: Ne152) were for Plato 'ideal' examples of inborn knowledge of ideal things.

Scientists have at least indirectly been influenced by Plato's ideal world. Natural science, as we now understand it, is, among other important things, a search for fundamental natural laws which are abstract and which objective reality obeys. Using natural laws, science derives the observed behavior of the world. These laws are something like Plato's ideal objects. Stating this is not to credit Plato with the parenthood for natural science; before Plato, the Presocratics were already searching for the underlying principles of nature. However, Plato's personal influence in his own lifetime and his eloquent writings through the course of history have emphasized the idea of the search for underlying principles.

It should be noted that Platonic ideas in a more direct form have had an influence on certain scientists in history: Kepler being an example par excellence of a Platonist scientist (Ca44). There is another point to note: the idea of fundamental natural laws is not a obvious idea to come to. Societies can exist, such as the Mesopotamia society of course, which see the willfulness of the gods in everything. And yet another point. It is not obvious that science can reach high levels in the absence of the idea of fundamental natural laws. It has been argued that the lack of such an idea was a component in preventing the development of Western-style science in China by Needham (see Co453–455).

If one turns to Plato's precise influence on astronomy in particular the situation is ambiguous. Letting him speak for himself—in voice of Socrates—Plato (in the *Republic*) had this to say:

> Thus we must pursue astronomy in the same way as geometry, dealing with its fundamental questions. But what is seen in the Heavens must be ignored if we truly want to have our share in astronomy ... Although celestial phenomena must be regarded as the most

> beautiful and perfect of that which exists in the visible world (since they are formed of something visible), we must, nevertheless, consider them as far inferior to the true, that is to the motions ... really existing behind them. This can be seen by reason and thought, but not perceived with the eyes (quoted from Pe24).

At least two interpretations are possible of this passage (Pe24–25). First, that Plato was advocating a purely artificial astronomy based on inborn knowledge and reasoning that would not be compared to or depend on astronomical observations. This interpretation seems to be that of Neugebauer (Ne152), who thinks Plato had little influence on astronomy and that his advice would have been its ruination if followed. The second interpretation is that Plato hoped that an axiomatic deductive astronomy (modeled on geometry) would explain the particulars of observed astronomy.

Maybe a right interpretation of what Plato actual intended advice to astronomers cannot be found (Pe25). However, the second interpretation, if not applicable to the passage quoted, is consistent with what Plato is reported to have said by the late writer Simplikios (c. 6th century BC). This is that the celestial phenomena should be saved by reducing their irregular motions to a superposition of uniform circular motions (Pe28). If one takes uniform circular motion for celestial bodies as a physical theory to be accepted as long as aids in furthering understanding of reality and rejected when it fails in that function, then Plato's reputed challenge is consistent with the scientific method. But it seems clear that uniform circular motions took on the nature of an a priori truth. As the history of astronomy unfolded, it is clear that uniform circular motions at first aided astronomical development as a means of reducing complex planetary phenomena to reasonable predictability, but then became a straightjacket curbing the search for deeper insight. The supposed authoritative endorsement of uniform circular motions by Plato and others (e.g., Ptolemy who violated it Pe] , probably did act eventually as a hindrance to the advance of astronomy.

Whether or not Plato issued the challenge to save the phenomena or it arose in some other way, Greek mathematical astronomers took it up. The first of these was in fact an associate of Plato, Eudoxos of Cnidos (c. 408–355 BC).

Eudoxos was a mathematician and in 368 BC he amalgamated his school of mathematics with Plato's Academy (Pe335). Some of his mathematical results are presented in Euclid's Fifth and Sixth Books (Pe335). His own works on astronomy are lost. However, from later writers he is known to have constructed first a sophisticated geometrical astronomical model using geocentric spheres with uniform rotations (Pe63–70). The Earth is motionless and at the center. The celestial bodies reside on some of the spheres and are carried by them in uniform circular motions. The idea of the spheres of course goes back to the Pythagoreans (see § 4.4; Pe52). It is not known if he gave these spheres a physical reality as Aristotle would do later in his version of Eudoxos model (Pe69; § 4.6).

In the model, there are 3 spheres for Moon and Sun 4 for each of the 5 planets. Another sphere was used to hold fixed stars (i.e., it was the Celestial sphere itself); thus there are 27 spheres altogether. For each celestial body, the spheres are nested. The outermost sphere rotates daily around the Celestial Pole and, of course, provides the daily rotation of all celestial phenomena for the body as seen from the Earth. Moving inward, each succeeding sphere is attached to the proceeding sphere by an axis not coincident with the proceeding sphere's: the rotation about this axis provides a motion superimposed on all the motions of the outer spheres. Using this mechanism most of the observed celestial motions could be built up.

For a simplified example, consider the Sun and two spheres (Pe65). The first sphere provides the daily motion: its angular velocity is the same as that of the Celestial sphere. The second sphere is attached at the angle of the obliquity, whose modern value is 23.44° (App. B, Table B3). (Note the obliquity varies a bit over the course of millennia.) It rotates once per year relative to the first sphere. (Note the distinction between sidereal and tropical year was unknown before Hipparchos' discovery of the precession of the equinoxes.) Combined the motions of the two provide the mean motion of the Sun. Eudoxos used in fact 3 spheres for the Sun: his third sphere was used to account for a supposed variation of the Sun from the ecliptic which was not defined as the Sun's path until the time of Hipparchos it seems. Actually, some modification of Eudoxos Sun model was needed since he assumed the Sun's velocity relative to the Celestial sphere was a constant.

The Sun's motion is relatively simple. The Sun does not show retrograde motions and, in fact, from a geocentric view its orbit around the Earth is nearly circular with a fairly constant velocity. The eccentricity of the Sun's orbit (the Earth's orbit from the heliocentric point of view) is only 0.0167 (Li14-3). However, the Earth does show some slight variation in angular velocity. This variation can be simulated by using more superimposed spheres. This is in fact is what a Kallippos of Cyzicus (c. 370–300 BC) did: he added two more spheres to the Sun model (Pe69, 321). Kallipos was a pupil of Polemarchos of Cyzicus (c. 340), pupil of Eudoxos (Pe381; No79).

One can see the beauty and the vileness of uniform circular motions. One can simulate planetary appearances with them, and some of those appearances do represent nearly uniform circular motion. But as one trys to represent ever more detailed observations, uniform circular motions can become an adjusted decomposition that do not yield any more information than is fed into them. Moreover, there will often be no uniquely good way to do the decomposition. To express this in the words a modern scientist would use: if you are given enough free parameters, any theory will work. The theory of uniform circular motions allows as many uniform circular motions as one wants: i.e., as many free parameters as one wants.

This critique of uniform circular motions is not a denigration of Eudoxos. He made a pioneering attempt at geometrical model of the planetary system (the universe as he understood it) that would explain all observables. His mechanism for explaining retrograde motions deserves praise, even though it was a complete deadend.

Eudoxos' retrograde motion device was the hippopede (horseshoe-curve) which looks like a figure-of-eight pasted on a spherical surface. The curve can be constructed from two spheres in uniform motion of equal speed: one sphere carries the other. Let the carried sphere be sphere 1 and the carrying sphere be sphere 2 Put a dot representing planet on sphere 1. If the sphere 1 is coaxial with sphere 2 and rotating oppositely, then the dot, which is moving in a uniform circle relative to sphere 2, is at rest in the frame of reference of a motionless outside observer. If the sphere 1 is tilted off the axis of sphere 2, the dot starts executing a figure-of-eight motion on an at rest spherical surface coincident with sphere 1. It is not easy to see that this should be so, but it can be proven by the classical methods of geometry (No71–76) or by modern spherical trigonometry. When the tilt gets to 90°, the hippopede passes through the poles of sphere 2.

If one forgets spheres 1 and 2 now, and just accepts that one has a planet tracing out a hippopede, one can attach the hippopede to the equator of another sphere: sphere 3. The hippopede is attached with bows along the equator of sphere 3. Sphere 3 has the Earth at its center and is rotating eastward with uniform circular motion and sidereal period of the planet. When the motion on the hippopede is in the same direction as the motion of sphere 3, the planet is in prograde motion. When the motion on the hippopede is opposite the direction of the motion of sphere 3, then retrograde motion is possible if the hippopede motion is great enough. Thus retrograde motion can be achieved. A fourth sphere carries sphere 3 in a uniform motion westward with the a period of one day. Thus we see how four spheres can account qualitatively for all the observed planetary motions that Eudoxos was aware of.

Qualitatively, but not quantitatively. As North discusses (No77) discusses Eudoxos' model does not have enough freedom to account quantitatively even to low accuracy for the motions of Mercury, Venus, and Mars. Another objection which disqualifies all variations on the Eudoxian scheme is that it does not account for the variation in the apparent sizes of the Moon and the Sun. This variation was discovered by Polemarchos of Cyzicos (c. 340) (Pe69). Autolycos of Pitane (c. 310) pointed out that this variation implying a varying distance could not be accommodated in a model where all celestial bodies moved on spheres concentric to the Earth. In concentric sphere models all celestial bodies are at fixed distance from the Earth always. The planets cannot be resolved into disks with the naked eye, and so could be assigned apparent sizes. But they do vary significantly in luminosity which strongly suggests significant distance variation as well.

************************* Still to be revised below?????

All though in principle Eudoxos model would have yielded quantitative predictions of reasonable ancient accuracy (perhaps a 5°????) for some objects there is no evidence the that all the parameters (rotation periods and angles) were ever inserted and predictions made. Adding more spheres to account for the motions left out by Eudoxos would have improved the accuracy at the cost of complication. However, mathematical astronomers after the time of Aristotle, at least, lost interest in Eudoxos' model. There are two main discrepancies from observations that could not be corrected.

First, in Eudoxos' model celestial bodies are set at constant (but unknown) distances from the Earth. Already at about the time of Eudoxos it was known that the Sun and Moon have variable angular diameter and therefore variable distance (without a crazy hypothesis of varying size). The varying brightness of the planets (which are not resolved) suggested that they too have a variable distances. Second, the retrograde motions which were reproduced by Eudoxos' model did not have the right shape against the Celestial sphere and were invariant for each planet: real retrograde motions vary from one synodic period to the next. The epicyclic model which develop in the 3th century BC at least offered the possibility of getting the varying distances and brightnesses right and the shapes of the retrograde motions.

One feature of Eudoxos' model must be emphasized. There is no scale. The celestial bodies and spheres could be put at any distance from the Earth. One just had to change the physical velocities of the spheres in proportion to their distances and the appearances stayed the same. The Greeks argued that the celestial bodies that moved faster across the sky must be closer if the physical velocities did not very greatly????. This is a good physical argument although not true. However, it did yield roughly the correct order of the celestial bodies from the Earth: Moon, closest, Mecury, Venus, the Sun next in some order, and the superior planets in their true order. Real distances, even relative distances were mostly beyond the Greeks' observational capabilities. Only for the Moon were they able to obtain a reliable distance (§ 4.7). For the Sun, their determinations were wildly inconsistent (§ 4.7). Probably there were no attempts to obtain planetary distances; if there were they could not have been accurate given ancient observing techniques. The epicyclic theory also gave no theoretical distances, not even relative ones.

4.6 Aristotle

Aristotle of Stagira (384–322 BC) bulks large in history. He was a pupil and then a critic of Plato, a teacher to Alexander the Great, and the founder of second great Athenian school, the Lyceum (or the Peripatetic) in 334 BC. He created a grand synthesis of human knowledge through writings and propositions on a wide range of topics: philosophy (speaking generally), ethics, logic, epistemology, physics (not quite as we understand the term today), metaphysics, mathematics, biology, politics, and literary criticism. In many of these fields his works would be held by many to be definitive, or at least the essential, for centuries (until at least c. 1600) in the Greco-Roman, Islamic, and European cultures. That Aristotelianism for most of that time appears to us retrospectively and to some in the 16th and 17th centuries as a stagnation does not invalidate Aristotle's achievements. Rather it demonstrates the persuasiveness and common sense of much that he produced. And it must be said that to an educated individual living in times when scientific progress was not apparent, Aristotlianism presented what appeared to be a philosophically sound, complete system that did not contradict ordinary observations (at least not flagrantly). Hence its long attraction.

Although a pupil of Plato, Aristotle broke from the master giving reality back to the world that we see and finding in experience of that world the only source of knowledge. However, we cannot deduce from this that Aristotle practiced the scientific method in an explicit manner. He made close and systematic observations of nature, particularly in his specialty biology. Although he must have experimented (at least in the realm of dissection????), he did not emphasize a need for continuous experimentation, quantitative measurement, and testing of theories under ever more extreme conditions. These are hallmarks of the modern scientific method. Certainly, the Aristotelians of Medieval and Renaissance Europe never felt their doctrine enjoined experimentation.

Another deficiency (relative the scientific method) was the lack of mathematization is Aristotle's procedure. Lack is not total omission. Aristotle was also a mathematician and was conversant with mathematical astronomy of his day (i.e., that of Eudoxos, whom he may have known). However, in the physics of motion he inhabited "the world of more-or-less (Co510)" rather than the "the Universe of precision (Co510)." There was no mathematical analysis in his physics. His views were common sense, or rather common experience, systematized. For example, aside from natural motions, nothing moved without a continuous force and to all motion there was a resistance (Pe105–

110). This explains common appearances, but we know it has no predictive power. Another example, heavy bodies fall faster than light bodies. This is commonly, but not always, observed and is caused mainly by differences in dynamic air resistance that correlate often with the density of the objects. A brick falls faster than a feather. Any deeper analysis that took into account different densities and shapes was not undertaken. To Aristotle the observed differences in falling speeds made the critical qualitative distinction. This example shows Aristotle in a particularly weak case. But his physics had other weaknesses. Even in Antiquity and the 14th century, there were critiques of his physics by his followers (Johannes Philoponos [c. first half of 6th century AD], and Jean Buridan [c. 1290–1360] and Nicole Oresme [c. 1320–1382]).

Aristotle's cosmology was consistent with common sense which is not necessarily a criticism. The Earth was a sphere and at rest in the center of spherically-symmetric, finite, eternal cosmos whose limit was the Celestial sphere. Beyond the Celestial sphere was nothing, not even empty space. For sphericity he had three valid arguments (which may not have been original to him of course). As one moves north and south the altitudes of the celestial bodies and the stars change in a manner consistent with a spherical shape. The shadow of the Earth on the Moon during a lunar eclipse is round no matter where the Sun is exactly. At great distances at sea masts can be seen when hulls are beneath the horizon.

For geocentrism his arguments were persuasive if not valid. First, if the Earth were moving relative to the fixed stars (and necessarily then not at the center of the cosmos), stellar parallax would be observed unless the stars were immensely remote or moving in fixed relation to us. From any perspective, the theory of remote stars is a fantastic one; it is just, as we know, a true one (or rather an adequate one). Aristotle notion of causality was teleological: things happen or exist for a purpose. A huge nearly empty cosmos seemed a useless absurdity????. The idea that the stars moved in fixed relation to the Earth could be rejected as ad hoc: i.e., an implausible (absurd!) fix-up.

The second argument for a resting Earth was that we do not detect that the Earth is in motion. For Aristotle, as mentioned above, there are natural motions and forced motions. On Earth, the natural motions are upward for the elements with levity (fire and air) and downward for elements with gravity (earth and water). Any sideways motion would require a force and would be felt or noticed in its absence: for example, by being swept off our feet as the world turns. These objections to a moving Earth are eliminated by the concept of inertia, but that would not appear until Galileo (although not in its modern Newtonian formulation).

The above arguments suffice to show that the Earth is at rest, but not that it is at the center. Aristotle's very plausible divorce of sublunary world and the heavens

4.7 Aristarchos of Samos and Heliocentrism

Aristarchos of Samos (c. 4th–3rd century BC, best guess c. 310–230 BC: Wa215) is as obscure, as Aristotle is well known: a dark star in the firmament. He is reported to have been associated with Museum of Alexandria (Wa201). Ptolemy reports that Hipparchos reports that Aristarchos made an observation of the summer solstice in 280 BC (To139). This is almost certainly the Aristarchos, not another Greek of the same name, and so gives the only firm date for Aristarchos' flourishing. There is one extant treatise, *On the Sizes and Distances of the Sun and Moon*, by Aristarchos. Unlike many other ancient works in the Aristarchos story, there is no question of attribution with this treatise: he is almost certainly proven to be the *Sand-Reckoner* author who refers to results in the treatise and assigns them to Aristarchos (Wa209; see below): almost certainly the Aristarchos and not another Greek of the same name. There are a few other references to Aristarchos from Antiquity which show him to have a considerable reputation as a mathematician and inventor: e.g., a reference from Vitruvius (quoted by Wa214).

On the Sizes is a work in theoretical astronomy. It is formulated in the axiomatic deductive manner, just what Euclid and Plato would have approved of. The method is geometrically ingenious, but the results are poor by modern standards: Aristarchos' input data was not good. Nevertheless, his estimates for the sizes and distances were better than most others of his day by being larger (Wa210). Later Greek astronomers got better results, but for the Sun they never achieved good accuracy: see Table 4.1 for a comparison of some of Aristarchos' results with other Greek astronomers and modern values. Better observational data would have helped Aristarchos in the case of the Moon, but for Sun because of its great distance (and small parallax) other techniques were needed????.

Table 4.1. The mean distances and parallaxes of the Moon and Sun

Source	D_{Mo} $\left(R_{\oplus}^{\text{Eq}}\right)$	$\theta_{\text{Par(Mo)}}$	$D_{\odot}$ $\left(R_{\oplus}^{\text{Eq}}\right)$	$\theta_{\text{Par}(\odot)}$
Aristarchos of Samos (4th–3rd century BC)	19	3°	360	9′30″
Hipparchos of Nicaea (c. 190–120 BC)	$67\frac{1}{3}$	0°51′	1245	1′23″
Poseidonios of Rhodes (c. 135–50 BC)	53	1°05′	13100	0′16″
Ptolemy (c. 100–175 AD)	59	0°58′	1206	2′51″
Modern	60.2684	0°57′2.60″	23454.78	8.794148″

Note.—The names and dates are drawn from Wa215, Pe346, Pe381, and To1, respectively. The ancient data is adapted from Pe48, except for the Ptolemy's solar distance and parallax. Ptolemy's solar parallax is taken from the *Almagest* (To265) and it is then used to calculate Ptolemy's solar distance. The precise modern values are referenced in Appendix B, Table B3.

Solar system parallaxes measured from the Earth are now always understood to be measured from a baseline of the Earth's equatorial radius; this is absolutely to be distinguished from stellar parallaxes which are measured from the baseline of the Earth-Sun distance (No103). The modern value for the Earth's equatorial radius is given by $R_{\oplus}^{\text{Eq}} = 6378.138\,\text{km}$ (App. B, Table B3). The distance d to any body in the solar system is exactly related to the parallax θ_{Par} by

$$d = \frac{R_{\oplus}^{\text{Eq}}}{\sin\theta_{\text{Par}}} \approx \frac{R_{\oplus}^{\text{Eq}}}{\sin\theta_{\text{Par}}} \ . \tag{4.1}$$

Ptolemy's solar parallax is about 20 times too large, but gained widespread currency. Tycho Brahe, unfortunately, assumed it, and so contaminated much of his solar analysis which otherwise used data much more accurate (Th226ff).

The most famous reference to Aristarchos and the herald of his glory is in a treatise called the *Sand-Reckoner* reputedly, but not certainly, by Archimedes (Wa204). The reference is just a digression:

> Aristarchus of Samos brought out *graphai* consisting of certain hypotheses, wherein it appears, as a consequence of the assumptions made, that the universe is many times greater than the 'universe' just mentioned. His hypotheses are that the fixed stars and the sun remain unmoved, that the earth revolves about the sun in the circumference of a circle, the sun lying in the middle of the orbit, and that the sphere of the fixed stars, situated about the same center as the sun, is so great that the circle in which he supposes the earth to revolve bears such a proportion to the distance of the fixed stars as the center of a sphere to its suface (quoted from Gi64).

This passage shows that Aristarchos at least proposed the heliocentric theory for the Earth with static Sun and stars, and that he understood that the stars had to be immeasurably remote (by ancient standards) to explain the non-observation of stellar parallax. Since the whole beauty of a heliocentric model is the order it makes among the planets (with the Earth one of them), it is almost certain that he made the other planets orbit the Sun too.

Aristarchos' *On the Sizes* is geocentric (No85). So we do not learn of his helocentrism from himself. There are only 5 other brief references to Aristarchos' heliocentrism in ancient works (Wa205-206): two by Plutarch (of Chaironea; c. 46–120 AD: Pe380), and the others by Diogenes Laërtes (probably 3rd century AD), Aëtios of Antiocheia (c. 100 AD: Pe300), and Sextus Empiricus (2nd century AD: Co297). The attribution of the works to the specified authors is at least somewhat suspect in all these cases (Wa205) just as is the Archimedes attribution. However, even if a work is wrongly attributed this does not imply, of course, that the citations of Aristarchos are invalid. The 5 other references, however, give little more information and may be mostly dependent on the *Sand-Reckoner* itself or some other common source (Wa205). Plutarch in one reference tells us that Aristarchos held heliocentrism only as a hypothesis, but that Seleucus of Seleuceia (or Babylon) (c. 150 BC: Pe391; No86) held it as decided opinion (Wa205). Seleucus is little known, but seems a significant thinker: he is reputed to have discovered the relation of the Moon to the tides and to have advocated the infinity of the universe (Pe391; No86). In the other Plutarch reference we find:

> Only do not, my good fellow, enter an action against me for impiety in the style of Cleanthes, who thought it was the duty of the Greeks to indict Aristarchos of Samos on the charge of impiety for putting in motion the Hearth of the Universe, this being the effect of his attempt to save the phenomena by supposing the heaven to remain at rest, and the earth to revolve in an oblique circle, while it rotates, at the same time, about its own axis (quoted from He108).

Cleanthes of Assos (c. 331-232 BC) was a Stoic philosopher and the successor to Zeno, the founder of Stoicism (Pe323). Diogenes Laërtes lists an *Against Aristarchos* among Cleanthes' works (Wa205–206).

Ptolemy does not name any heliocentrists, but refers to "certain people" who make the Earth rotate on its axis and make Earth and heaven have any motion at all (To44). He concedes that there may be nothing against these ideas in the observed of the celestial appearances, but objects that physically the idea of a moving Earth is ridiculous (To45). Ptolemy's physics is essentially Aristotelian with Stoic leanings (Pe77–78). Aristarchos is probably among these "certain people": Seleucus maybe as well. Others might include Hicetas of Syracuse (5th century BC), a Pythagorean astronomer, who assumed an axial rotation, but in the Philolaic system (Pe54): how this is possible without getting rid of the Central Fire and Anti-Earth I do not know. Nothing much else is known of Hicetas (Pe346). Ecphantos of Syracuse (c. 400 BC), also a Pythagorean, possibly a disciple of Hicetas, did dispose of the Central Fire and Anti-Earth, and put the Earth in the center of the cosmos with an axial rotation to explain the daily westward motion of the heavens (Pe54, 331). Heracleides of Pontos (c. 390–310 BC), a disciple of Plato, believed in the axial rotation of the Earth (Pe54, 342). The notion that Heracleides propounded a geo-heliocentric theory, putting Venus and Mercury in orbit about the Sun (which is still orbiting the Earth) may be just due to a misreading of a text (Pe54): if so, it seems to have been an early misreading and thus an effective accidental innovation (Pa118; Pe54–54).

We are left with Aristarchos and Seleucus as the only true heliocentrists of Antiquity. Since they are evidenced to be both deep thinkers, one can reasonably suppose that they understood the essential advantages of heliocentrism: (1) the order of the planets from the Sun can be derived (Ku174), (2) the relative distances of the planets from the Sun can likewise be derived (Ku174), (3) the physical distinction between superior and inferior planets disappears (Ku172–173), and (4) retrograde motions are simply explained (Ku171). These advantages were realized by Copernicus and were his essential reasons for turning to heliocentrism (Ku171) despite his essentially Aristotelian views (Pe77–78; Ku147). It is possible that Ptolemy perceived these advantages since he concedes the celestial appearances could be saved perhaps with a moving Earth (To45). Ptolemy could have constructed epicyclic heliocentric models just as he did geocentric ones. But he had a strong Aristotelian bent, and had not, perhaps could not, become frustrated with the endless variations on Ptolemaic systems that the coming centuries were to reveal to be possible. Fourteen centuries later, Copernicus felt the full force of that frustration and turned to heliocentrism (Ro4, 21). Ptolemy remains Ptolemy; Copernicus, Copernicus.

But what of Aristachos? What were the *graphai* in the quotation from the *Sand-Reckoner*. Heath (He106) translates the word as book, but *graphai* has more than one meaning: its first meaning seems be 'drawings' (Gi65). Other meanings are 'outline', 'figures', or 'writing' (which could be in a book) (Wa203). No other evidence for a book exists. Wall suggests (with some reason) that the *Sand-Reckoner* author may indeed having been referring to a book: *On the Sizes* (Wa210).

The idea is that the *Sand-Reckoner* author, intentionally vague, attributed Aristarchos' certain heliocentrism to the geocentric work to give it a respectable, but undefined, reference. Alternatively, the *Sand-Reckoner* author may just have been confused on the correct reference and left it vague therefore, intentionally or not. In any case, the evidence for a real Aristachan book on heliocentrism is slight.

Gingerich (Gi65) suggests a minimalist interpretation of Aristarchos' heliocentric advocacy. The heliocentric idea was discussed by Aristarchos, perhaps in Alexandria, possibly with Archimedes (who was a student there) present, and some illustrative drawings (i.e., those *graphai*) were shown. And there was no book. The animus of Cleanthes, however, suggests that the heliocentrism was talked of for a time among philosophers. So maybe there was a book. Sans new evidence coming to light there is no way to decide. However, there is some point in speculating maximally for a moment.

Let us say that Aristarchos was struck by the advantages of heliocentism. The very fact that Aristarchos had poor data available to him may have caused him to focus more on the inherent simplicity of the heliocenctric idea than worry about the complications that would be needed in either a predictive Ptolemaic or Copernican system. Moreover, even with his poor determinations, Aristarchos knew that the Sun was at least ~ 2 times the diameter of Earth: maybe he argued physically that the biggest body in the cosmos should be the central object. Going further, maybe he wrote a book, even a big book. Maybe he could have applied the Eudoxian concentric spheres to the Sun to "save the phenomena": they would have worked better without the need for hippopede for retrograde motion. If he had done so much, why was it not better known; why did Plutarch say that Aristarchos held the heliocentric idea only as a hypothesis; why did Ptolemy not say more. Maybe Aristarchos feared persecution from the Cleanthoi of the world or maybe just ridicule. Maybe he was just secretive; a Pythagorean, who did not wish to divulge more of the truth to the ignorant mass. And so this imagined book was seen by few and vanished.

Obviously, I have made a fantasy. But Copernicus held the heliocentric hypothesis in his thirties at least (Gi68; Ko143, 148), distributed one manuscript giving an outline of his early ideas (the *Commentariolus*), worked out his ideas in great detail, and was sitting on them still when he was in his late sixties. Only the importunities of his disciple Rheticus and his great friend Tiedeman Giese managed to free *De Revolutionibus Orbium Coelestium* from Copernicus' reticence to publish (Ko160–161, 167). Without such friends, Copernicanism would probably have gone by a different, later name. Maximal as well as minimal explanations are possible. The minimal may be more sensible so as not to waste energy and to create imaginary problems (which is the sense of Ockham's razor), but the potential richness of actual, if unknown events, should not be forgotten either.

Even without a book, only a bruited about idea, why did so little come of heliocentrism: it was near to vanishing without a trace. We have already discussed Ptolemy's rejection. His Aristotelianism was too strong, and perhaps the full force of heliocentrism did not strike him: it being an already old and bypassed idea in his day (c. 150 AD). And there were few others. After Aristarchos, only three great creative theoretical astronomers turned up in Antiquity: Apollonios (c. 262–190 BC: Pe307), Hipparchos (c. 190–120 BC), and Ptolemy again. None of these became an Aristachan. Without a brilliant exponent, an idea so far from Aristotle's philosophical soundness, so far from the center of the Earth, was perhaps doomed to near oblivion.

A last bit of if-history to close this section. Say Aristarchos had forcefully presented his ideas in a book, famous or infamous. Would this event alone have substantially rewritten history? Could heliocentrism have become a powerful ancient heresy? Could it have triumphed sooner? much sooner? Could a Scientific Revolution have been precipitated in Antiquity by an Aristarchan revolution just as at least partially one was by the Copernican revolution?

I suspect the answer to the last question is no. My argument (synthesized derivative argument) for the decline of Greek science is given § 4.12. Changing one theory only to a better one, even a very significantly better one, even one that demanded a breach in Aristotelian physics, probably would not have prevented the decline of Greek science. After all, the Greeks had the theory that the Earth was a small sphere in a large space. They found good proofs for this theory too. But though marvelously thought-provoking, the spherical Earth in a huge space did not start a relentless quest for natural laws. It did not even send Greeks and Romans off on exploring expeditions to verify the total sphericity and to discover the whole Earth. As for the breach in Aristotelian physics, one can imagine a qualitative principle of planetwide, but not small-scale, inertia or adhesion filling the gap. Copernicus was may have been thinking along such lines of such a principle (Ro16). I can see no compulsion to Scientific Revolution in heliocentrism.

The answers to the other if-history questions are maybes. There is no reason that the heliocentrism could not have batted around more than it did. Perhaps, it could have found acceptance as a calculational device when armored with epicycles and deferents to make it predictive. After all, dyed-in-the-wool Aristotelians would only accept the Ptolemaic system in the calculational device guise (Pe69). Then later heliocentricism could have become a dogma itself: one that just happened to be right. It seems to me, that a minor bit of historical re-engineering could have saved a millennium or so, at least in the narrow field of mathematical astronomy.

4.8 Apollonios of Perge

Apollonios of Perge (c. 262–190: Pe307) is one of the greatest of ancient mathematicians. He is said to have studied with the successors of Euclid in Alexandria (Pe307), but that it has been questioned if he was there long (No89). He developed the theory of conic sections in his treatise *Conica* and we owe to him the terms ellipse, parabola, and hyperbola (Pe307). His mathematical reputation rests solidly on his extant works; his astronomical reputation is great too, but known only through references. The astrologer Vettius Valens (c. 160 AD) reports seeing tables of the Moon and Sun by Apollonios, but these may have been by a another Greek of the same name. Ptolemy in the Almagest (c. 150 AD: To1) reports on Apollonios development of epicycle and deferent planetary theories (To555ff). It seems

4.9 The Flowering of Hellenistic Science

the foreword:

> Seeing moreover in you, as I say, an earnest student, a man of considerable eminence in philosophy, and an admirer [of mathematical inquiry], I thought fit to write out for you and explain in detail in the same book the peculiarity of a certain method, by which it will be possible for you to get a start to enable you to investigate some of the problems in mathematics by means of mechanics . . . for certain things first became clear to me by a mechanical method although they had to be demonstrated by geometry afterwards because their investigation by the said method did not furnish an actual demonstration. But it is of course easier, when we have previously acquired, by the method, some knowledge of the questions, to supply the proof than it is to find it without any previous knowledge (quoted from Pe97).

4.10 Hipparchos

Hipparchos of Nicaea (c. 190–120 BC) is generally acknowledged to be the one of the two greatest astronomers of Greek Antiquity: Ptolemy being the other. Hipparchos was certainly the greater observer of the two and perhaps the only Greek astronomer to see and act on the need for accumulation of large quantities of accurate astronomical data.

A source for this data that he utilized was Mesopotamian astronomy; Hipparchos may be the principle conduit for Mesopotamian data and conventions into Greek astronomy and hence to posterity. He had eclipse records for 600 years in the past that could only have come from Mesopotamia (Pe50). He used the Mesopotamian division of the circle into 360° and sexagesimal arcminutes and arcseconds. (An slightly earlier appearance of the Mesopotamian 360° circle and sexagesimal arithematic in Greek is in the *Anaphorics* of Hypsicles [No93]). It is a pity that the sexagesimal system was never or at least almost never used consistently by the Greeks and their successors (Ne22); Copernicus, however, was more nearly consistent than many others (Ne22).

In theoretical astronomy, however, Hipparchos did not take over the linear algebraic methods of the Mesopotamians (although he should have known something of them), but built on the work of Apollonios (Pe73). He did not attempt planetary theory (as far as it known), but only theories for the Sun and Moon). He created a Sun model using a simple eccentric circle, and fed in the observationally determined parameters which Apollonios did not do so far as we know (Pe73). This model was quite successful for prediction and was adopted by Ptolemy (Pe78). It must be noted that the Sun is the simplest case. Geometrically, it really can be considered as orbiting the Earth in an ellipse with eccentricity $\epsilon = 0.017$. To first order in small ϵ, an eccentric circle of radius a and an ellipse of semi-major axis a are identical (see App. B, Table B7). Thus an eccentric circle model for the Sun should be excellent.

For Moon, Hipparchos used a deferent circle centered on the Earth and an epicycle with retrograde rotation (i.e., clockwise as viewed from the north) which is opposite the deferent's prograde rotation (counterclockwise as viewed from the north). This model was less successful, but Hipparchos was probably able to predict the possibilities of solar and lunar eclipses using it. It must be remarked that the Moon is a more difficult case than the Sun. The lunar orbit is more eccentric ($\epsilon = 0.055$) and the orbit itself has a retrograde rotation (the regression of the nodes) with a period of 6796 days (18.61 mean Gregorian years). And there are other subtler effects These and the regression of the nodes are are due to perturbations by the Sun and other planets: the Earth-Moon system is only very crudely a two-body system. Not until Laplace (1749–1827) were Moon's motions brought under good control (Pa305). (Perfect lunar prediction is in fact ruled out by the minute Chandler wobble: an erratic orbit of radius $\sim 15\,\mathrm{m}$ (period 13 to 14 months) of the physical Earth about the rotation axis probably due to major earthquakes [Pa368ff and Sm77ff].)

Hipparchos has a number of great astronomical achievements (Pe49–50). According to Pliny he observed a new star (*stella nova*) which may have been a nova (in the modern sense of the word), supernova, or just a variable star. Possibly this new star is the one recorded by Chinese observers in 134 BC. Hipparchos was supposedly inspired by this observation to create a catalog of fixed stars. The important fixed stars were well known to be invariable, but a catalog would reveal if other stars varied or even came and went. This catalog contained $\sim$ 850 stars (Pa129) in a variety of coordinate systems. This catalog was lost, but it is believed to have formed the basis of Ptolemy's catalog of 1022 stars (Pe81). Since only a few thousand stars are visible to the naked eye (Pa444), Hipparchos' catalog can be considered very complete. (Friedrich Argelander at Bonn created an atlas [Uranometria Nova] of all naked eye stars [presumably only for a mid-northern latitude????]) that included only 3256 objects: stars, variable stars, and nebulae [Pa444].) Hipparchos' catalog was not the first in history, however. The 4th century BC Chinese astronomer, Shih-shen is said to have compiled a catalog of 809 stars in at least 122 constellations (Pa88 and Pa90). Only fragments have been preserved.

It is conjectured that Hipparchos invented the magnitude scale (a brightness scale) and used it in his catalog; magnitudes certainly precede Ptolemy's use of them in the *Almagest* (To16). The magnitudes ran from 1st magnitude (brightest stars) to 6th magnitude (faintest observable stars). Naturally this magnitude scale was qualitative and based on intercomparison. However, it would have a long history; it was taken up by Ptolemy and passed onto posterity. Actually the eye can make finer judgments than 1 magnitude and Ptolemy would qualify his magnitudes with the words 'greater' and 'smaller'.

After the invention of the telescope fainter (i.e., higher) magnitudes would be added to the scale???? (Pa445). In the 19th century, it was noted that the eye responds approximately logarithmically to physical brightness (i.e., intensity or energy flux). In 1850, N. R. Pogson at Oxford proposed that the magnitude scale be regularized according to the equation

$$M = -2.5 \log b + C \ , \tag{4.2}$$

where M is magnitude, b is brightness, and C is a constant chosen as appropriate (and kept secret if at all possible) (Pa446). With this equation magnitudes are logarithms of the approximate base 2.512 and 5 magnitudes correspond exactly to a brightess difference of 100. Magnitudes are nowadays measured in reasonably well-defined wavelength bands: e.g., the B (blue), V (visual: actually yellow), and R (red) bands. Magnitudes can now, of course, be calculated for any brightness and so can run from negative to positive infinity. It has to be remarked that it is a perennial irritation having a logarithmic measure that runs the wrong way (i.e., brighter is smaller, fainter is larger) and that has a completely stultifying base.

Hipparchos' single greatest discovery was of the precession of the equinoxes (Pe49). This slow periodicity could only be noticed because he had centuries of reasonably accurate Mesopotamian data. The precession of the equinoxes is better described as a westward (i.e., retrograde or clockwise) recession of the Earth's axis and thus the Celestial Pole about the ecliptic pole (i.e, a line perpendicular to the ecliptic plane). The axis maintains its (mean) obliquity of $\sim 23.5°$. Traced on the sky from below, the circuit of the axis is counterclockwise. Hipparchos discovered the precession by noting that the sidereal year was slightly longer than the time between the vernal equinoxes (i.e., the tropical year) (Pe49). The modern values for the sidereal and tropical year are 365.25636 days and 365.24220 days, respectively. The precession causes the equinoxes to move westward along the Zodiac: hence the traditional name: precession of the equinoxes. Hipparchos determined only a lower limit of 36″ per year (or 1° per century). This value, however, became accepted as the best value in Antiquity. A modern precise value is 1.3969712° per Julian century (Table B3). The Islamic astronomers al-Battānī (c. 850–929) and Ibn Yūnus (c. 940–1009)) established quite good values of 54″ per year and 51″ per year, respectively (Pe162ff). However, another one, Thābit ibn Qurra (c. 826–901) introduced the idea of a variable precession (called the trepidation) based on overconfidence in ancient observations. This phantom, the trepidation, which had Copernicus had accounted (Pa197), was exorcised by Tycho Brahe, who found a constant precession of 51″ per year (Pa215).

The physical cause of the precession was first explained by Newton using his laws of motion and gravitation. The Earth is actually an oblate sphere (polar radius 6357 km and equatorial radius 6378 km). This asymmetry causes other celestial bodies to exert unbalanced torques on the Earth which on the average are perpendicular to the angular momentum vector of the Earth's spin. The torques try to rotate the Earth perpendicularly to the Earth's axis: the result is a precession just like a toy top. The Moon and Sun (with the Moon dominating) are the causes of the torque (Sm75). There is also a planetary contribution to the precession which is about 1/40 of the total due to the Moon and Sun and opposes their action. The planetary contribution is due to the motion of the ecliptic, however, not the motion of the Celestial Pole (Mo72).

The total period of precession is 25780 years (Ze11). A consequence of the precession is that Polaris is only temporarily the pole star. About 3000 BC the pole was near Thuban in Draco and in 14000 AD, pole will be near Vega. Currently, Polaris is within 1° of the pole and will get closer to it until 2100 AD and then move away (Se29). (To be precise, the declination of Polaris in year 2000 will be 89°15′51″ [Ho20].)

Hipparchos was also interested in determining geographical latitudes and longitudes (which had come in too use even before 300 BC) by astronomical means (Pe49ff). Latitude is straighforward: a gnomon and the declination of the Sun suffice. Longitudes are more difficult. Determining them from the days of travel between two location was very imprecise. In order to make better determinations, synchronized clocks and transits of the meridian can be used. The time between transits in two locations translates directly into difference in longitude: e.g., if a transit occurs at place X 2 hours after occurring at place Y, then X is 30° west of Y. The difficulty in ancient times was in synchronizing clocks. Celestial phenomena themselves were the usual clocks and they obviously provide only a local time (i.e., celestial time). Mechanical clocks would only be invented in the Middle Ages, and sand and water clocks do not transport well. Hipparchos suggested using entrances (or exits) of the Moon into the Sun's shadow during a lunar eclipse as synchronization signal. All people on the night side of the Earth can see a lunar eclipse (unless clouded out) and, of course, they all see the entrance at the same time. Lunar eclipses, however, are sufficiently rare and their exact time of occurrence difficult for the ancients to predict. It does not seem that Hipparchos's idea came into practical use. (Ptolemy used the fact that lunar eclipses occur at different celestial times as one of his arguments for the sphericity of the Earth [To40]).

Hipparchos determined distances to the Sun and Moon (see Sect. 4.7, Table 3.1). The Moon distance is fairly accurate, but the Sun distance, although much better than Aristarchos', is much too small. Hipparchos' method for deriving the Moon distance is quite clever (Pa129). Consider the

Moon halfway entered into a lunar eclipse. Let angles α and β be the solar and lunar parallaxes, respectively: the line tangent to the Sun and Earth and passes through the center of the Moon is perpendicular to radius of the Earth that subtends the parallaxes. Let angles δ and κ be the angles at the Earth's center subtended by the Sun and Moon, respectively. Geometry shows that

$$\alpha + \beta + \pi = \delta + \kappa + \pi \tag{4.3}$$

or

$$\alpha + \beta = \delta + \kappa \ . \tag{4.4}$$

Now α is the much the smallest angle and can be neglected. Thus the lunar parallax is

$$\beta \approx \delta + \kappa \ . \tag{4.5}$$

From the lunar parallax β the distance to the Moon follows from trigonometry.

A final possible contribution. Astrological literature often quotes Hipparchos (Ne187). Given Hipparchos' close connection to Mesopotamian sources and the probable origin of Hellenistic astrology in the 2nd century BC, it is at least plausible that Hipparchos was a contributor to astrology.

4.11 Ptolemy

Klaudios Ptolemaios (Κλαύδιος Πτολεμαῖος), or, modern version, Ptolemy of Alexandria (c. 100–175 AD: To1) ranks as the foremost theoretical Greek astronomer. Little is known about his life. Presumably he worked at the Museum of Alexandria or at least had access to its library. His astronomical works are addressed to (i.e., dedicated to) an unknown Syrus. One of his teachers was probably Theon, who is not to be confused with the 4th century Theon of Alexandria. Judging from Ptolemy's first name, he was probably a Roman citizen. (Roman citizenship was still a civil distinction until Emperor Caracalla made all subjects, excepts slaves and the *dediticii* [whoever they were], citizens possibly in 214 [Chambers 1966, p. 38].) His second name, and the evidence of his writings, shows him to have been of Greek (and maybe Macedonian) cultural heritage. In Medieval Islamic and European times, he was identified with Ptolemaic dynasty and sometimes illustrated wearing a crown (Pe80; No106). Although worthy of a crown, Ptolemy was not likely scion of that royal line (except possibly in remote way that everyone has royal ancestry): Ptolemy was a common name in Greco-Roman Egypt (No106).

Ptolemy wrote a number of works, some of which are lost. The important survivors include the *Optics*, the *Geography*, *Planetary Hypotheses*, *Handy Tables* (for astronomical computations), and the *Almagest* (Pe76–77). The series seems to make up an encyclopedia of applied mathematics: but it was also a series of original and outstanding work too. The *Almagest* is the greatest as its name implies. It is work on theoretical astronomy. Called by Ptolemy Mathematical Systematic Treatise (Μαθηματικὴ Σύντξις), it probably acquired the Greek nickname greatest treatise (*μηγίστη* Σύντξις) (To1–2). This mutated into the Arabic *Al-majastī*, and thence to the Latin *Almagestum* and the modern *Almagest*.

The *Almagest* is a work in thirteen books (in the classical meaning of the term). Toomer's modern translation spans 613 pages (To34–647): a largish modern treatise. It probably took years in the writing and a lifetime of scholarship. It may have been finished about 150 AD and is probably almost the earliest of Ptolemy's surviving works (To1). In the *Almagest*, Ptolemy takes up the astronomical research of Apollonios and Hipparchos apparently just where they left off 4 and 3 centuries earlier. He is our own witness for this since he cites and describes their work (Pe382–383; No106). In fact, without Ptolemy we would know little about their achievements (No89; Pa125). There seems to have been no great advances in the 3 centuries between Hipparchos and Ptolemy (No119), although astronomical and, especially, astrological work continued in between them (No119ff; Ne187). This stagnation in science is a subject we will return to in § 4.12.

In Book I gives Ptolemy's philosophical and physical assumptions. These are essentially Aristotelian with a Stoic influence (Pe78). An absolute Aristotelian, of course, would not deviate into epicycles (as Ptolemy does) from Aristotle's geocentric spheres (Pe69). Also in Book I he gives some mathematical chapters and a table of chords which serve him in place of modern trigonometric functions (To57–60). The Ptolemaic chord function is related to the sine function by

$$\mathrm{ch}(\theta) = 2R\sin\left(\frac{\theta}{2}\right) \tag{4.6}$$

where 'ch' stands for chord function and $R = 60$. In Book II, Ptolemy develops spherical trigonometry (a great original contribution) starting from the Menelaos theorem of plane trigonometry (Pe78). The astronomer and mathematician Menelaos of Alexandria (c. 100 AD) seems to be the only near contemporary whose theoretical work Ptolemy relied on.

Books IV–VI give Ptolemy's solar and lunar theories (Pe78-80). The solar theory is Hipparchos' unchanged: this a simple simple eccentric circle model. The fact that the Sun orbits the Earth (in a geocentric view) in an almost unperturbed ellipse with an eccentricity of only 0.0167 makes an eccentric circle model a good approximation. The eccentric circle and ellipse orbits agree to first order in small eccentricity ϵ (see App. A2).

The reason for near ellipticity of the Sun's orbit is the weakness of all multi-body effects. For a two-body system of spherically symmetric masses interacting only under gravitational force Newtonian physics (which is very nearly correct in the weak solar system gravitational fields) dictates perfect elliptical orbits of the two bodies around their common center of mass (Shu463–466). The motion is in an absolute inertial frame of Newtonian physics which for the solar system is effectively the frame of the fixed stars. Each body can be regarded as orbiting the other in an ellipse of the same shape with a different size scale: the relative sizes of these various ellipses are determined by the relative masses. If one body is overwhelmingly dominant in mass, the elliptical orbits of the second body relative to the first body and to the fixed stars are virtually identical. The situation just described basically applies to the Sun-Earth system. The gravitational effects of the other planets and the Moon are weak perturbations of the Sun-Earth interaction. When perturbations are not weak, orbits will be distorted from perfect ellipse and will often tend to rotate relative to the fixed stars: i.e., the orbit shape will rotate.

Ptolemy's solar model does not correspond to well to a modern eccentric circular model. His solar eccentricity was 0.042 (Ne192; Pe74). Such a large discrepancy from the modern value may be partially due to poor data, but is likely partially because Ptolemy's model gives the Sun a uniform spatial motion. The real motion is non-uniform whether an elliptical model or the analog eccentric circular model is used. The non-uniformity in motion probably got absorbed into Ptolemy's (and Hipparchos') eccentricity parameter????.

For the lunar model, no simple model can be accurate. For the Sun-Earth system, one can ignore the Moon as a weak perturbation. But for the Earth-Moon system, the Sun's perturbative effect cannot be ignored. The Moon's orbit (which is nearly elliptical with eccentricity 0.54900489) rotates westward relative to the fixed stars with a period of 18.61 Jyr (Mo731): thus 19.34°/Jyr. Trying to deal with the Moon's eccentricity, the non-uniform motion around its elliptical orbit, the rotation of its orbit, and still other lesser perturbative effects using a superposition of circular motions is tough challenge. The Moon's orbit is also tilted with respect to the ecliptic plane by a mean??? angle of 5.14°: this causes the Moon to move in ecliptic latitude. Ptolemy's lunar model is of the deferent-epicycle kind. However, to account for all the Moon's anomalies (i.e., differences from uniform circular motion), he put the made the deferent center off the Earth and made it rotate on a circle too. The deferent center and the lunar epicycle rotate westward (No112; Pe79). The deferent itself rotates eastward and accounts for the main lunar motion.

The lunar latitude variation was achieved by tilting the deferent orbit plane from the ecliptic plane by 5° (No112). The orbits of the planets were titled too (with some extra complications for the inferior planets) to account for their variations in ecliptic longitude. The epicycle orbits were given different inclinations from the ecliptic than the deferent inclinations. In the later *Handy Tables*, Ptolemy made the epicycle planes parallel to the ecliptic plane. North describes the reasoning for this improvement (No118).

Ptolemy's lunar model gives a good account of the lunar longitude and latitude motions. But there is a major descrepancy from observations that Ptolemy was surely aware of, but does not comment on (No113). In his lunar model the Moon's distance varies between $33R_\oplus$ and $64R_\oplus$ (Pe80). The angular diameter of the Moon varies linearly with the lunar distance. Thus Ptolemy's model predicts that the Moon's angular diameter should vary from $\sim 0.5°$ to $\sim 1°$. Such a variation would be readily apparent even to casual eye observation. From Ptolemy's *Planetary Hypotheses* we know that he took his models as serious attempts to model the physical world, not merely as calculating devices (No112-113). Thus he would have had to have judged his lunar model as badly flawed. He would have known that the Moon's real motion in space was probably best only remotely like his model's motion.

In the case of the Moon, this reality check was available to Ptolemy: we easily resolve the Moon's disk and even the ancients could measure the distance to the Moon accurately. Ptolemy himself gave the Moon's mean distance as $59R_\oplus$ (see Table 4.1). This compares well with the modern mean distance of $60.2683R_\oplus^{\rm Eq}$: the distance varies up and down from this value by 5.5 % for a total range in variation of 11 %. For the Sun there was a less blatant discrepancy. Ptolemy's eccentricity of 0.042 would lead variatian in distance and angular diameter of 8 %. Ancient solar distances were wildly inaccurate (see Table 4.1), as Ptolemy may well have known, and so solar distances provided no check on his solar model. But the Sun's disk is resolved, of course, and the variation in its angular diameter had been discovered by Polemarchos of Cyzicos (c. 340 BC). The total variation is only 3.3 % or

Books VII–VIII give a star catalog with (ecliptic) longitude, latitude, and magnitude.

To summarize it briefly summarize the Almagest: It contains an

The practical ending of creative Greek astronomy with Ptolemy finds its particular explanation in the explanation for the general decline of Greek science, which is the subject of § 4.12.

4.12 Whatever Happened to Greek Science Anyway?

It is generally agreed among historians that Greek science, after a period of continuous progress, declined well before the end of the Greco-Roman civilization (Co???). The timing and reasons for decline are in dispute. However, I give an analysis. Nothing in this analysis is original, except in that it is my own synthesis of the ideas of historians of science. The analysis is based mostly on the analyses of Cohen (Co???) and the historians he discusses, particularly Joseph Ben-David and G. E. R. Lloyd.

Greek natural philosophy (which encorporates science in my view) began with first Presocratic, Thales, circa 600 BC. The next two hundred years were the age of the Presocratics. In this time, each generation brought forth new thinkers who responded in some innovative manner to the ideas of their predecessors or contemporaries. The 5th century BC was also the time of Hippocrates of Cos (2nd half of 5th century BC) and his followers, and their innovations in medicine (Pe348). In the 4th century BC, there was Plato, Aristotle, Epicuros, Zeno (the founder of Stoicism) and host of other names. In particular, the 4th century BC saw the rise of axio-deductive geometry (Ne???) This was an outstanding development in itself and as a model for how science should be done; geometry, however, is not an empirical science and the axio-deductive procedure is only part of the modern scientific The 3rd century BC, had the founding of the Museum of Alexandria and a group of great names, more or less associated with it: Euclid, Aristarchos, Archimedes, Eratosthenes, Apollonios, and Ktesibios. Up to this point, there seems to have been no generational break and innovation seemed continuous. But after circa 200 BC, there appears to have been a change.

In the 2nd century BC the only great name is Hipparchos. In the period 100 BC–100 AD, there may not to be any great names at all: it is a question of definition of course. Poseidonious the astronomer, Philo and Hero the polytechnologist, and Meneloas??? were at least notable. In the 2nd century AD, there were two great names Ptolemy and the physician Galen of Pergamon (Pe337). Both men wrote great books that would become standard sources until circa 1600. However, they seem to have been rather lonely figures without great (at least great and remembered) contemporaries. In the case of Ptolemy, the work that he builds on is mainly that of Apollonios and Hipparchos along with some theorems of Meneloas???? His teachers were primarily books. The books were there: anyone could have taken them up in the two centuries since Hipparchos's death and made a new contribution, but seemingly no one did.

After Ptolemy there is not much to add. Pappos of Alexandria (c. 300) made important mathematical contributions (Pe374). In the 5th century, Johannes Philoponos (c. 500–550) made some original contributions to physics in the form of a commentary to Aristotle (Pe359). Other noted names seem to have been mostly important only as commentators: Theon of Alexandria (c. ????), Hypatia of Alexandria (c. ????, daughter of Theon), both astronomers and mathematicians; Proclos of Lycia (c. 415–485), a neoplatonic philosopher at the new??? Academy; Simplikios of Cilicia (6th century). Simplikios was among the last of the neoplatonic philosophers (Pe391). In 529, Emperor Justinian closed??? the new?? Academy: it was a pagan institution in a state now almost wholely Christianized; its closure can be seem as a symbolic closing of Greco-Roman Antiquity. Simplikios and others from the Academy left for Persia. Simplikios' commentaries on Aristotle were influential in the Medieval Islamic and western European societies.

In this resume, I put the decline at 200 BC. Others may put a different date: Lloyd (Co253) puts it after Ptolemy and Galen. However, I put the decline as starting when I perceive a first discontinuity between great names. This dating is consistent with the Ben-David model of science in 'traditional' societies.

The Ben-David model (Co254) postulates that discontinuity is natural in traditional societies. These societies are essentially all societies before the Scientific Revolution which began in circa 1600 in Europe. In traditional societies, there can be a scientific innovator now and then or a string of such innovators, but inevitably there is a fading of innovation, followed by a period of no innovation and maybe even loss of knowledge. The reason for the lack of continuous cumulative science is the weakness of support for science. Not that support is altogether absent (though it may be), but just that it is insufficient. I will take 'support' to mean a broad range of things, both external and internal to science. In the following I discuss under a number of headings what I see as the support weaknesses of Greek science. A general discussion of Ben-David model is beyond my scope.

4.12.1 The Pay-off

A key support for modern science is the economic and spiritual support given to modern science in exchange for its contributions to the economic and physical well-being of society. Science is seen as a golden-egg-laying goose; a good investment. Obviously, a most visible support for science is the money that pays numerous scientists and supplies their research needs. But there is also the spiritual support that scientists feel in doing work that is recognized and honored for its tangible benefits.

But the pragmatic reasons for supporting science were fairly negligible prior to 1600. By and large, new technology did not result from the investigations of natural philosophers and scientists. The Roman engineers who built the roads, aqueducts, the Pantheon (with its concrete dome???) did not so far as we know rely on the results of natural philosophers. Perhaps they could have learnt something Archimedes about statics, but they may already have known his results empirically. And it is empirically that they worked, not searching for universal laws. The same applies to the sailing ship builders and the farmers. Land surveyors needed only relatively simple geometry, not the extensive proofs, theorems, and lemmas of Euclid. Certainly ancient business accounting and all numerical work would have been improved by using a place value notation with an explicit zero such as the astronomers (e.g., Ptolemy) used although inconsistently (Ne11,22). Evidently, nonscientific numerical work was insufficiently arduous to require such an innovation.

The above statements cannot be pressed too dogmatically. The polytechnologists of Alexandria, Ktesibios (c. 200 BC), Philo (c. 1st century AD) and Hero (c. 1st century AD), and, of course Archimedes all made practical inventions. However, it is not clear how much application their work found????. It is plausible that the war engines Archimedes invented and used in the siege of Syracuse (Flachière & Chambry 1966, p. 209) found a place in the arsenal of the Roman legions????. Certainly, medicine must is a field where the benefits of the scientific approach were recognized: hence the widespread respect for work of the Hippocratean tradition and Galen.

Astronomy qua astronomy had no practical applications. And ancient navigators did not need it. They relied on simple astronomical observations for directions and tried to stay in sight of land which is fairly easy in the narrow seas of the Mediterranean????. Some astronomy was needed for astrology (if that can be considered a golden egg), but most astrologers did not need to observe the stars or understand the deeper waters of mathematical astronomy. All a run-of-the-mill astrologer needs is a few astronomical tables of modest accuracy and some imagination. The work of Hipparchos and Ptolemy was more than sufficient for these: no further developments were needed. Nevertheless for some sublime spirits such as Ptolemy himself with his *Tetrabiblos* and in later times Tycho Brahe (???) and Johannes Kepler (???) astrology was an important study. For them at least, improving astrology was an inspiration for improving astronomy. But their ambitions were superfluous for common astrologers and clients.

The lack of practical interest in science has several possible reasons. The main one may be that science was just insufficiently advanced to help technology in many cases. Even after 1600, technological applications of science only slowly developed (Co???). Another may be an attitude of disdain for practical application on the part of natural philosophers. Plutarch attributes such a disdain to Archimedes and Plato (Flachière & Chambry 1966, p. 208–209). It is possible that the attitude is stronger in Plutarch himself than in the people he attributes it too (Pe92). However, it is true that no writings of Archimedes on his practical inventions has come down to us???. It is also true that his works on physics are based on the axio-deductive method with postulates and derivations with no references to experimental work???. This does suggest a bias against writing about practical work.

Yet another reason may have been the lack of perception that science could yield practical benefits. In the 16th century, Francis Bacon and René Descartes (Co???) both proclaimed the practical utility of science. Their words testify to their insight more than to actual events in their time (Co???). In his time, Archimedes' famous words 'Give me a place to stand, and I shall move the earth' (taken from Pe308) did not become a motto and a metaphor.

There was one field of research that did promise practical benefits: alchemy (Pe141ff). The essential goal of alchemy was the making of gold by what we now would now call chemical processing. Alchemy arose in Alexandria sometime before 100 AD, perhaps as early as 200 BC. Most of the earliest writings are attributed, super-implausibly, to mythical or famous historical persons: e.g., Hermes, Moses, and Cleopatra. From the Greco-Roman world it spread to the Islamic civilization, and from there to western Europe where it still found adherents in the 18th century, notably Isaac Newton. Obviously alchemy failed of its quest. However, along its path many basic chemical procedures were discovered and many chemical apparatuses were invented. The theories of alchemy tended to be mystical and Aristotelian, and were connected to astrology. With such a framework, it is not surprising that theoretical chemistry did not make much headway among the alchemists.

Although the ends of alchemy were eminently practical and it founded an enduring tradition, it is not easy to see it as attracting popular support because it shunned that. Alchemy was a gnosis, an arcane science for an inner circle of adepts. The language of its texts is coded and opaque and sometimes defies interpretation. Nevertheless, alchemy's golden promise probably ensured its survival when the more lucid and rational parts of natural philosophy were neglected.

4.12.2 Where was Technology?

In the last subsection, it was argued that science was weak because it supplied little to technology. Here the argument is almost reversed: science was weak because technology was weak. If one looks at late Medieval times through the Renaissance, there was a striking progress in technology in western Europe. Here only a few of the most obvious items will be mentioned. Circa 1300 there was the introduction of the mechanical clock, convex lenses, and spectacles????. About 1450, moveable type printing and concave lenses (a boon for myopiacs) appeared. Throughout the 15th and 16th centuries there were improvements in ships, maps, and navigational techniques. It should be noted that not all these technological innovations originated in Europe

The Renaissance-to-Enlightenment scientists worked against a background of obvious technological progress. In a vague and general sense, it is plausible that a general sense of innovation in technology was a spur to do likewise in science. Additionally, the precision of the newer technology was perhaps a challenge to greater precision in the description of nature: i.e., to go to a mathematical description of nature. Against these vague notions, some harder evidence can be atested for the stimulation of science by technology. The case par excellence is the invention of the telescope. The telescope was invented in 1608 in the Netherlands: the exact history of the discovery is clouded. Within a year Galileo and others were uncovering a host of never-seen-before celestial phenomena.

Another facet of a world of technological progress was the fact that innovations could find an application and reward. This opened to at least some sages the possibility of a dividend from their studies. Galileo (a prototypical figure) was quick to exploit his improved telescopes (much better over the earliest crude examples) for worldly gain. Not by sale—that would have been *déclassè*—but by strategic gifts to princes.

In contrast, the Greco-Roman world showed a rather leisurely pace of technological progress. Building and civil engineering improved noticeably. But overall progress must have been sufficiently slow that a person would not perceive it in their lifetime and a perception of cumulative development may have been lacking. The introduction of the water wheel in late Antiquity is a rare example of a notable innovation that provoked some comment????.

The historian of science Benjamin Farrington??? proposed that the slave labor economy of the Greco-Roman world militated against the occurrence of a Scientific Revolution. The cheapness of impressed human labor forestalled the need for labor-saving technology and thus reduced the need for science to supply that technology. This thesis may not be totally supportable in a direct sense. But it is at least plausible that technological ingenuity was blunted by the slavery, and that blunting (as this subsection argues) blunted science. It is also possible that the pervasive hypocrisy of the Greco-Roman elite about slavery (good for others, not for themselves) contributed to the possibly significant attitude of disdain for practical application of science as unworthy of a philosopher (see § 4.12.1).

4.12.3 Besides the Tangible

Besides the tangible there is the intangible. Science answers questions about the natural world. Curiosity about the natural world is natural. Ergo science will be supported. The logic is not flawless. Certainly, curiosity (and even stronger words) are an ultimate drive for science and natural philosophy: it is almost undebatably that scientists and their patrons are driven by curiosity. But not all people will feel the urge to go beyond what is already available: the mythology or the established natural philosophy. Thus, the works of the natural philosophers of Antiquity were recopied again and again, and many of the most important ones were passed onto posterity through many generations who did not add to their contents. Thus, mathematics, astronomy, and natural philosophy were part of the ancient educational curriculum, but people were taught the subjects, not taught to pursue them. The Eastern Roman Empire (the Byzantine Empire) verifies this: a thousand year history with scarcely an original natural philosopher to name despite possessing much of the literature of Greek Antiquity (Pe151). Parenthetically, it needs to be said that Byzantine civilization was far from sterile overall: in art, literature, theology, theological controversy, politics, military activity, Byzantium has its achievements (e.g., Vryonis 1966; Hoxie 1966).

Is an explanation needed for science not being all things to all humans? Better to just give the explanation and avoid yet another indifferent digression. Science does not provide a meaning for life or say how to live. Those are the provinces of religion or philosophy in the broad sense. The great philosophers of the 4th century BC—Plato, Aristotle, Epicuros, Zeno the Stoic—had long and widespread vogues because of their overall philosophical systems that grappled with meaning-of-life and how-to-live. Natural philosophy was only a component of their systems. For many of their followers the natural philosophy component needed no amendments. However, it is fair to say that the consolation of philosophy was not enough for most. The mainstream of humanity finds its meaning-of-life and how-to-live in schools other than the Academy.

4.12.4 No Printing Then

There was no printing then. Indulging in if-history, one can ask what if the ancient Greeks had had printing (and paper, of course, to print on) would that have saved science from declining. Fortunately, comparative history can be done. There are two effectively independent examples of printing civilizations: Europe after 1450 (Gies241) and China after circa 700 (Ts1).

The effect of moveable type printing in Europe was certainly impressive. The availability and affordability of books increased dramatically. The scholar, student, ordinary person could keep a small library for private edification, something scarcely possible before????. The works of the Greco-Roman and Islamic world (usually in Latin translation) could be given a wide diffusion: thus ancient and Muslim learning was revealed and revealed not to be final. People could rise rapidly to fame and influence just through their printed words: e.g., Erasmus and Galileo. All classes of people, even of very low status could benefit. And all classes of people, in particularly the lay class, were exposed to a host of competing ideas and innovations. Naturally, censorship became a necessity. But it could be evaded: a determined author could publish his scurrilous (or otherwise unwelcome) work in one or other of the multitude of independent jurisdictions. It has been claimed (Eisenstein 1979; see also Co359) and questioned (Co357–367) that printing was essential for the Scientific Revolution. Certainly, it is hard to imagine the course of a Scientific Revolution sans printing, but that does not to rule out its possibility. A stronger claim is that printing was a sufficient cause for a Scientific Revolution. Cohen argues for the negative based on the effect of printing in China (Co366, 476).

My remarks and conclusions about Chinese printing are largely derived from those of Tsien (Ts). In China, block printing (i.e., printing where a page is set in a solid block of wood or metal) was in use $\sim$ 750 years before moveable type printing appeared in Europe. The greatest demand for the earliest printing sprung largely from the desire by devotees of Buddhism to massively reproduced religious works (Ts9–10). It came into large scale use circa 1000 (Ts369) and it did have profound effect: how could it not have. Learning was certainly more widely spread; works were saved for posterity; life was enhanced. But the overall effect seems to have been more stabilizing than mutagenic (Ts382). Almost all printing after circa 1000 by government, private printers, or businesses was dominated by works of Confucian scholarship (Ts378). Although government printing was not dominant in quantity????, given the hierarchical nature of China the government line must have set the pattern. Radical deviations from the Chinese norm were not introduced. Was a Scientific Revolution a radical deviation? I would not think so if it came about in the right way: whatever that way may be. However, no Scientific Revolution occurred. Given Chinese society, printing even on large scale was evidently not a sufficient cause.

By way of digression, a word should be said on the possible transmission of printing from China to Europe. It seems likely that block printing (first appearing in Europe in the 14th century???) was a transmission from China. The techniques of block printing first used in Europe are nearly identical to those used in China. The agents of transmission could well have been the European travellers who visited China in the 13th and 14th century: Marco Polo was not alone; Christian missionaries, in particular, were active in China????.

The transmission of moveable type printing is less certain. China had moveable type printing since circa 1000, but did not use it extensively. The nature of Chinese script with thousands???? of characters made moveable type printing rather unmanageable (Ts8). Since most Chinese printing was reprinting????, it was more sensible to produce the blocks for books which could be kept and re-used repeatedly. There is no known chain connecting to Johann Gutenberg to China and on grounds of technique???? no evident copying on his part of Chinese or other eastern models. However, he may well have learnt something through travellers and merchants. On the other hand an independent invention is also plausible. There can be no conclusion without more evidence.

4.12.5 Early Christianity and Science

Christianity became the dominant religion of the Greco-Roman world in the period $\sim$ 300–400. This is indisputably after the decline of Greek science. Thus Christianity was not cause of the decline. Still it is worth asking if Christianity had a positive or negative affect on natural philosophy in late Antiquity. The conclusion that I have come to is that it was more or less neutral.

The Christian leaders and writers (the Christian Fathers of the Church) showed both hostility and respect for the pagan philosophers. The hostility because the pagan philosophers were not of their faith and were rivals, and because their writings would lead people astray. The respect for at least some of their achievements, their virtues, and the truth to be found in their works. A quote from St. Augustine (354–430, bishop of Hippo in North Africa) illustrates this:

> There is no reason why we should not have studied literature just because pagans say that Mercury is its god, nor need we avoid justice and virtue because they have dedicated temples to justice and virtue and preferred to worship in stone what should be borne in one's heart. Rather, every good and true Christian should understand that truth, wherever found, comes from his God (quoted from Levine 1966, p. 222–223).

Naturally, the hostility and respect could not be equally balanced overall or in any individual. A general conclusion about the interaction of early Christianity and Greek philosophy is well out my range of scholarship.

In regard to natural philosophy, it does not seem to have occupied the Christian Fathers too much. Their attitude was often that natural philosophy was a vain philosophy and unnecessary for salvation in any case???? (e.g., St. Ambrose of Milan [c. 340?–397] [Dr212]). Such an attitude could not be expected to lead to a scientific renaissance, but would not likely oppose interest in many areas of natural philosophy. One area that become controverted was cosmology.

The Bible makes a number of cosmological statements that treated literally or mystically, even when they seem most like simple analogies, led to cosmologies radically at odds with Greek cosmology (No226ff, Dr207ff). For example, from Isaiah XI, 22: "it is He . . . that stretcheth out the heavens as a curtain and spreadeth them out as a tent to dwell in." This passage led to a doctrine of the heavens as a tent or, more Biblically, a tabernacle in the writings of the Syrian bishop Severianus (c. 400) and other Patristic writings (Dre211–212). (Patristic means pertaining to the Christian Fathers or their writings).

The most egregious Patristic writer is probably Lactantius (c. 240–320), who was a professor of rhetoric and tutor to Crispus, the son of Constantine. He had considerable classical learning and wrote with a Ciceronian style (Levine 1966, p. 217). Classical knowledge notwithstanding, in his *Divine Institutions* he ridicules and refutes the spherical theory of the Earth and the existence of the existence of the Antipodes. (Geometrically, antipodes are diametrically opposite points on a sphere. The Antipodes loosely seems to mean the southern hemisphere of the Earth????. The word antipodes comes from the Greek *antípous*: with feet opposite.)

The notion of the Antipodes is, of course, necessary with the spherical theory of the Earth possibly first introduced by the Pythagoreans. The Greeks had shown that the part of the Earth they were familiar with was consistent with a sphere and certainly could not be flat. Astronomical considerations made idea of completely spherical Earth entirely plausible. But, of course, the Antipodes were a theoretical extrapolation that would not be demonstrated until the time of the oceanic voyages of the 14th and 15th centuries. It seems that Antipodes were first introduced by the Pythagoreans as a necessary consequence of their spherical Earth (Dre37; Pe18). The Pythagoreans supposed that the Antipodes were inhabited. At some point the idea arose that the equatorial zone of the Earth was impassible due to extreme heat (Dre220) or that the Antipodes were otherwise unreachable (Dre213). Consequently, the inhabitants of the Antipodes could not be descended from Adam: a thought which distressed some of the Fathers (No226). Lactantius thought it absurd to believe that at the Antipodes people and all other things exist hanging upside down (Dre209).

Although ideas similar to those of Lactantius and Severianus were spread and adopted by some Patristic writers, they were not universally held and they were not carried on with interminably. The writers and bishops, St. Ambrose of Milan (340?–397), St. Augustine, and Isodore of Seville (c. 560–635) do not take a firm position and so were probably open to believing in Greek cosmology (Dre212, 213, 220). Their caution was necessary since deviations or perceived deviations could have called down fulminations on their heads. The Venerable Bede of Jarrow in England (673–735) wrote without quibbling about the spherical shape of Earth and the cause of the variation of the day with latitude and season (No228). By the time of Gerbert of Aurillac (c. 950–1003), who was Pope Sylvester II (999-1003) and an introducer of Islamic science into Europe, the flat Earth and tabernacle heaven theories had lost most credibility with Christian scholars (Pe340; Dre226).

4.12.6 Conclusion

My own view, and it seems unexceptional, is that knowledge of many sorts including especially science and technology tends to be cumulative in human society as a whole. The special mention of science and technology is because they are based on the nature of the natural world: e.g., the law of gravitation is the law of gravitation in any society; the steam engine is the steam engine in any society. There can be periods of stagnation, even periods of loss of knowledge: the most dramatic of these being labeled dark ages. But because of the broad geographic spread and multitudinousness of human cultures, such accidents of history as stagnations or dark ages have been local affairs and they did not halt the general progress science and technology. Given human nature and the nature of reality that allows tremendous material and spiritual rewards to be reaped from the application of the modern scientific method, it seems to me to have been highly probable (but probably not inevitable) that a Scientific Revolution would occur some place, some time in human history. Accidents of history determined that the Scientific Revolution (circa 1600) and later the industrial revolution occurred in Europe where there was a super-favorable set of circumstances. The capabilities produced in part by these revolution allowed them to be spread world-wide, forestalling any independent recurrence.

Before Europe circa 1600, there were other occasions when an approach to a Scientific Revolution occurred. The florescence of Greek science is an often cited example. However, the revolution did not then ensue perhaps because of the unfavorable circumstances discussed above. Maybe just a few more geniuses (a Galileo, a Kepler) and a few more technological breakthroughs (the printing press, the telescope) would have ignited the revolution. But then again maybe not.

5. Islamic Astronomy

5.1 Prelude: Indian Astronomy

By comparison to the Greeks earlier and the Islamic astronomers later, Indian astronomy in the period c. 300–800 BC cannot be said to have made outstanding progress per se. Nevertheless, in this period (and not at any other time) Indian astronomy seems to have been the world leader. Its greatest achievements were mathematical methods (which are also, of course, very important for things other than astronomy) and in just keeping the practice of astronomy flourishing so that it could be spread westward again to the Islamic world.

The Indian Brahmagupta's Siddānta

5.2 Hindu-Arabic Numerals

The origin of Hindu-Arabic numerals is somewhat obscure. There are four elements to these numerals which certainly originated in diverse times and places: a decimal base, place value notation, an explicit zero symbol, and decimal fractions. The decimal base certainly originated independently all over the world: its commonness is just a result of anatomy as Aristotle noted (Bo3). Maybe the earliest occurrence of a written account of finger wisdom comes from Egyptian *Book of the Dead* Ne9. A dead king wishes to cross the river to the underworld. The ferryman queries him: "This august god (on the other side) will say, 'Did you bring me a man who cannot number his fingers?' (Ne9)." The king recites a rhyme numbering his fingers revealing himself to be a powerful magician and crosses to the land of the dead.

The use of decimal place value notation is found in China as early as the 13th century BC in the time of the Shang dynasty (Ron5). Chinese decimal fractions appear as early as 330 BC (Ron36). Zeros, however, were expressed by a vacant space until quite late. A circular zero symbol is found in print from 1247 AD, but may have been in use at least a century earlier.

The earliest report of Hindu decimal place value notation without a zero is from a Syrian Bishop, Severus Sebokt in 662 AD (Bo235). The first appearance of the Indian zero is from an inscription of circa 870 AD (Ron4; Bo235): just a dot or very small circle (Bo261). Decimal fractions do not seem to have been used (so far as I can tell) at least before Hindu notation got passed onto the Islamic world. Whether the Hindu decimal place value notation and zero appeared independently in India is impossible to say. The decimal place value notation may have come from China or south-east Asia where it is attested to from circa 600 AD.

An alternative suggested by Neugebauer (Ne189) is that the both the Hindu decimal place value notation and zero may have originated as an adaptation of the sexagesimal place value system (with an explicit zero) used in texts derived from the Greco-Roman world (Ne189). It seems that the original version *Paulisa Siddhānta* contained a decimal place value expression in words for 7800 in number words (Ne189). The Islamic astronomer al-Bīrūnī of Khwārizmı(c. 973–1050 [Pe317])(who wrote a book on Indian mathematics and culture, *India* [Bo263]) attributed this work to the Greco-Roman astrologer Paulus Alexandrinus of the 4th century AD: *Paulisa Siddhānta* probably dates from 4th century at the earliest (Ne175; see also No165).

Given that Mesopotamian sexagesimal system with a zero and sexagesimal fractions had been in existence since at least 300 BC, but without a 'sexagesimal point' (absolute size had to read from context) (Ne20, 27), it is perhaps a bit strange that an equivalent decimal system developed so slowly. Greco-Roman, Indian, and Islamic astronomers all used the Mesopotamian sexagesimal system (although not often consistently) (Ne22) and could have developed the decimal system quickly just by analogy. Maybe for their computations they were content with the sexagesimal system and for everyday use required nothing more than crude decimal and other systems.

The same book, Brahmagupta's Siddhānta which introduced astronomy to the Islamic world (see 5.2), also gave the Islamic world Hindu numerals (hereafter Hindu-Arabic numerals) where they found extensive use (Bo260). However, Islamic mathematicians and astronomers did not use them consistently (Ne24). The Hindu-Arabic numerals seem to be largely used only in mathematical work. Astronomical tables used alphabetic numerals. Greek and Coptic alphabetic numerals were used in Egypt for centuries after the Arabic conquest of 641–643 (La39). The numeral shapes evolved somewhat both in India and in the Islamic world. The western Arabic Gobar numerals were the ones that are the ancestors of the European and modern numerals (Bo261). Al-Kashi of Kashan (d. 1429 [Pe361]), who worked at Uleg Beg's observatory in Samarkand made significant use of decimal fractions and regarded himself as their inventor, although he may have learnt of them from China or elsewhere (Bo268).

The introduction of Hindu-Arabic numerals into Europe was a gradual process. Gerbert of Aurillac (950?–1003), Pope Sylvester II (999-1003), is first Christian author to describe the Hindu-Arabic numerals, but without the zero (Pe340; Bo276-277). Adelard of Bath (c. 1090–1150) (Pe299) and John of Seville in the 12th century had further explained the Hindu-Arabic numerals to western Europeans, but they still did not achieve wide attention (Bo278; Gies225). Later authors who popularized their use were Alexandre de Villedieu (fl. c. 1225), John Sacrobosco or John of Halifax (c. 1190–1236 [Pe360]), and Leonardo of Pisa, called Fibonacci (for son of Bonacci) (c. 1180–1250). The Hindu-Arabic numerals only slowly came into common use probably due to conservatism and because businessmen were afraid they could be fraudulently altered (Gies 226–227). Moreover, the abacus was widely used and advantages of decimal place value notation are much more apparent in pen-and-paper calculations. Late in the 14th century, however, they came into general business and everyday use (Gies226–227).

A few more words can be said to complete the story of numerals and simple arithmetic operations. Fibonacci book *Liber Abaci* (*Book of the Abacus*) appeared in 1202 (Bo280). Despite its title it was not on the abacus, but rather treated problems, algebra, and strongly advocated Hindu-Arabic numerals. Fibonacci's most famous problem is discussed in Appendix A, § A3. The Fibonacci sign for zero is '0' which in Arabic is *zephirum* (Bo280). *Zephirum* and variants give us 'zero', of course, but also 'cipher'. Fibonacci writes his algebraic equations in words; only in diagrams does he use symbols for unknowns (Gies226). Binary operation signs appear much later: minus and plus in the 15th century, equal in the 16th century, and division in the 17th century.

Decimal fractions using the equivalent of decimal point appear in Europe as early as 1492 (Bo307), but decimal point itself did not appear until the 1590s. Two astronomers, Giovanni Antonio Magini (1555–1617) (a Galileo rival) in 1592 and Christoph Clavius (1537–1612) in 1593 are both attributed with its introduction (Bo334). Scottish mathematician John Napier (1550–1617) made the decimal point popular more than 20 years later (Bo333–334).

5.3 The Setting of Islamic Science

Islamic astronomy is a part of Islamic science. The expressions, Islamic astronomy and Islamic science, are used here to mean astronomy and science in the Islamic world in the period circa 750–1450: this seems to be the accepted historiographical usage (Co384). Before circa 750, there was no significant science supported by Muslims in the still only recently born Islamic society. After circa 1450, Islamic science declined. When science revived in Islamic countries in the 19th and 20th centuries, it was modern science and thus part of a different history. In this section, we will sketch setting of Islamic science: viz., the general history that of Medieval Islam.

?????Muhammad (570?–632) was born in Mecca in the Hijaz in the eastern part of the Arabian peninsula. In about 610, he began his open career as the Prophet of Islam, the Way. In 622, he moved to the nearby city of Medina and became its religious and political chief. Both by word and sword, he spread Islam and its political hegemony. After his death, his immediate successors, the caliphs (kalifat Allah????, the deputies of God) proceeded likewise in a century of conquest with Arabian tribes as their armies. The Sasanian??? Persian empire was overthrown, the provinces of Syria and Egypt were taken from the Byzantine empire, and the conquest to the east extended to the Indus?????. Westward, all of north Africa and Visigothic kingdom of Spain were taken. In 732, Charles Martel (????) defeated the ????

5.4 Islamic Astronomy: Circa 760–1200

Al-Khwārizmī of Khwārizm (d. 860 [Pe362]) was Persian mathemetician and astronomer. He worked in Baghdad and may have been involved in the measurement of the Earth's circumference ordered by Caliph al-Ma'mūn.

5.5 The Marāgha School

The Marāgha School is a modern name given to a small group of astronomers who either worked at the Marāgha Observatory in north-western Persia (modern north-western Iran???) or made developments based on work of the astronomers in the first category. In fact there are only four names currently considered to be significant: Nasīr al-Dīn al-Tūsī (1201–1274 [Pe370]) Mu'ayyad al-Dīn al-Urdī (d. 1266 [Sa245]) Qutb al-Dīn al-Shīrāzī (1236–1311 [Pe385; Sa245]) and Ibn al-Shātir (1304–1375 [Pe350; Sa245]) The collective work of these four astronomers is now often considered to be the brightest achievement of theoretical astronomy in the Medieval Islamic civilization. It is a scientific tragedy that their work was almost unknown in Europe in before and during the Copernican revolution and was generally forgotten until the 1950's (Sa259). The significant exception to obliviousness in Renaissance Europe, the person who knew at least something significant of their significant work—indirectly to be sure—was the significant man himself, Copernicus.

How the Marāgha school came to be, what its work was, and how it was transmitted are the topics of the following subsections.

5.5.1 Hulagu Khan

In his lifetime Genghis Khan (1167–1227) conquered an empire beginning with his native Mongolia that extended over modern Mongolia and the north and west of modern China (Ru237ff). His grand strategy in essence was reward vassals, be terroristic to those who refused to be vassals and were insufficiently strong to resist, and expand continually. His army based on his Mongol tribes grew with every vassal made. Its essence was cavalry, archers and lancers, men hardened on desert treks and raids, discipline, and unheard of numbers (Ru242). His sons and grandsons continued his conquests and became rulers of sub-empires. As long as they stayed united and reinforced each other from their respective territories, they could bring overwhelming force to bear at any point and continue to expand. Once the descendents lost cohesion (essentially in the time of the great-grandsons), the Mongol Empire splintered, continuous expansion halted, and eventually collapses followed.

In 1251, Mongka Khan, a grandson of Genghis Khan, was elected supreme Khan: the fourth supreme Khan counting from Genghis Khan himself (Ru294ff). He re-animated the policy of conquest assigning the two main thrusts in China and Middle-East to his brothers Kubilai and Hulagu, respectively. Kubilai proceeded to conquer that part of China not already in Mongol hands with great skill and, insofar as a conqueror can be, was humane: he in fact converted to Buddhism. Hulagu, like Kubilai was more educated than most Mongol princes, and liked to have learned men in his service. He was interested in philosophy, alchemy, and astrology. Unlike Kubilai, Hulagu was savage to his enemies in the style of his grandfather.

Northern Persia and Azerbaijan had been annexed to the Mongol Empire by the Mongol general Chormaqan in 1231. In about 1253, the Mongol general Kitbuqa restored and extended Mongol authority in Persia in preparation for the main effort. This began when Hulagu brought a huge Mongol army across the Oxus in January 1256. By the end of 1257, he had defeated rulers of the western of Persia, the Assassins (as called by the Crusaders) and captured their capital Alamut. His next objective was Baghdad, which invested on January 22 1258.

Baghdad at that time was still the capitol of the Abbasid Caliphs. They had long been shorn of their great empire, but in 1258 were again masters of their surroundings at least. Despite rival caliphates that had existed, the Abbasids still had a strong theoretical claim to being the heirs to the authority of the Prophet. And Baghdad was still wealthy and populous city, a fabulous and fabled city. On February 10, Hulagu's army broke into the city, the Caliph was soon put to death after revealing his secret treasure hoard, within forty days tens of thousands of inhabitants were massacred. Because of the dead, there was fear of epidemic and Hulagu retired to Azerbaijan (modern north-west Iran) where he had a base established near Marāgha.

Hulagu extended his war to Syria and early in 1260 both Aleppo and Damascus had fallen: at Aleppo Hulagu was merciful; Damascus surrendered not counting on mercy. But the unity of the Khans was breaking. Hulagu's brother Ariqboga was elected supreme Khan in 1259 after Mongka death without Hulagu or Kubilai's approval and a civil war was in progress in Hulagu's rear between Ariqboga and Kubilai. Also Hulagu's Muslim cousins of the Golden Horde (as they called themselves [Ru294]), whose territory was on his northern flank, were disapproving of his anti-Muslim actions. Fearing troubles to the north and east, Hulagu withdrew many of his men soon after conqueroring Damascus. His general Kitbuqa remained in charge in Syria. In September of 1260, Kitbuqa was defeated and killed at Ain Jalud by the Mameluk army from Egypt. Western Syria was lost to the Mongols and they never recovered it or moved further to the south-west (Ru313).

Kubilai defeated Ariqboga in 1261, and thus restoring some of Hulagu support from the Far East. But the Muslim Mongols of the Golden Horde and Turkestan were closer and held him in check. Kubilai gave Hulagu the title Ilkhan (Khan of the west????) and hereditary government of south-west Asia: the Ilkhanate of Persia. Hulagu died in 1265 and was succeeded by his son Abaga. Hulagu's great grandson, Ghazzan acceded to power in 1295 and soon thereafter converted to Islam changing his title to sultan. The Mongol power in Persia evaporated after 1335.

5.5.2 Marāgha

Nasīr al-Dīn al-Tūsī, a Persian, was one of the great scholars of Islam steeped in Hellenistic philosophy and science (No193ff). He had been in the service of the Grand Master of the Assassins, but after Hulagu's conquest he enter the personal service of the Khan and was present at the conquest of Baghdad. Al-Tūsī was effectively Hulagu's chief astrologer (Sa38). He may have been one of the astrologers who tried to discourage Hulagu's attack on Baghdad (Ru302).

In 1259, al-Tūsī convinced Hulagu to establish an observatory at Marāgha with al-Tūsī himself as first director. Marāgha (modern Maragheh) is 80 km south of Tabriz. It is roughly equidistant from the Caucasus Mountains to the north, the Caspian Sea to the East, modern Turkey to the west, and modern Iraq to the south-west. The Marāgha Observatory had a large scientific library, a librarian, a staff of at least 10, and was equipped with expensive instruments: an armillary, a mural quadrant, parallactic rules, etc. There was at least one Chinese astronomer, Fao Mun-ji, making the observatory international as befits an institution of a supra-national empire. Actual observations at Marāgha began in 1262 and the observatory survived some manner until 1316 (Sa177). The earliest sources give the rational for the observatory as the need to obtain new planetary parameters to improve astrological prediction. This goal was achieved by the production of the Zīj-i Ilkhani (Ilkhanic Tables) probably created under al-Tūsī's supervision prior to his death in 1274 (Sa177–178, No193).

The great achievement of the Marāgha astronomers, however, was not in observation but in theory. Al-Tūsī himself wrote on logic, philosophy, theology, and mathematics. He was of the Aristotelian persuasion. In a popular work, *The Treasury of Astronomy* (*Tadhkira*) he continued the critique of Ptolemy that had begun with with al-Haytham most noticeably (Sa13).

Al-Tūsī's single most remarkable achievement was the discovery of a theorem known in modern times as the Tūsī couple theorem. The theorem states:

> If a first circle rolls without slipping inside the circumference of a second fixed circle with radius twice as great, then any 'physical' point attached to the rolling circle moves along a straight line (a diameter of the fixed circle) (adapted from No194).

The proof is straightforward. Consider the fixed circle as being centered on the origin: it has radius $2r$. Consider the rolling circle at time zero as having its center at the point $(r, 0)$. We assume that rolling circle center goes counterclockwise: its rim must go clockwise. The point where the two circles touch at time zero is $(2r, 0)$. Label the 'physical' point on the rolling circle point A. The coordinates of point A as a function of time, t, are given by:

$$[x(t), y(t)] = r\left[\cos\theta_1(t) + \cos\theta_2(t), \sin\theta_1(t) + \sin\theta_2(t)\right] \ . \tag{5.1}$$

The initial condition implies that $\theta_1(t) = \theta_2(t) = 2\pi n$ at time zero, where n is any integer. Since the rolling circle rolls without slipping the y coordinate must be zero at all times: thus

$$\theta_2(t) = -\theta_1(t) \ . \tag{5.2}$$

Using the facts that $\cos(x) = \cos(-x)$ and $\sin(x) = -\sin(-x)$, we obtain

$$[x(t), y(t)] = 2r\,[\cos\theta_1(t), 0] \ . \tag{5.3}$$

Thus point A moves along the x-axis only and for continuous rolling of the rolling circle will oscillate between $(2r, 0)$ and $(-2r, 0)$.

There is nothing special about point A. Thus any 'physical' point on the rolling circle will move along a diameter that is determined by the two points—the antipodes by gosh—on the fixed circle where that physical point will touch the fixed circle during a complete cycle. If the angular velocity of the rolling circle is a constant, then all the 'physical' points on its rim are executing simple harmonic motion.

The Tūsī couple (as the two circles are called) became a handy device for model planetary model construction in the Ptolemaic vein. One can create cyclic non-uniform linear motions from superimposed circles with only uniform motions.

A devil's advocate might give some credit to the Khan.

5.6 Al Andalus: Islamic and Christian Spain

5.7. The Course of Islamic Science

As in the case of Greek science, the decline of Islamic astronomy cannot isolated from the decline of Islamic science as a whole. However, even more so than in the case of Greek science, the discussion of the decline must be cautious. A great wealth of Medieval Islamic manuscripts have never been studied. The record to analyzed is much less complete than it could be.

The course of Islamic astronomy, except in an internalist view, cannot be studied apart from the history of Islamic science as whole. And the course of Islamic science itself cannot be studied by an internalist treatment only. That is to say, a treatment that only deals with the intellectual problems and solutions of science. Such internalist treatments are essential, but cannot explain why the scientific enterprise started, was sustained for awhile, and then declined. A full explanation of the course of Islamic science is well beyond our scope and as far as I can tell has not been done by anyone one yet (e.g., Co387, 409; Sa52). However, I have gleaned and synthesized some analysis from the literature.

In historiography, the term Islamic science refers to science done in the Islamic society in the period circa 750–1500 (Co384). After this period science became dormant (in the terminology I use) until it was re-stimulated by contacts with European science; this later science is, of course, part of international modern science, and cannot characterized as Islamic science in a historical sense.

The period starts with the great effort of translation to Arabic of Greek, Indian, and Persian of scientific works principally under the sponsorship of the 'Abbasid Caliphs in Baghdad (Co386; No174ff, 179; Sa52). For astronomy and mathematics, the most important early translation was of an India work in Sanskrit. According to tradition, this work was brought by Indian astronomer to Baghdad in the reign of Caliph al-Mansūr (r. 754–775 [La66]) in the 760s (Sa72). Only an abridged version was made, and only one text based on this work has survived. Thus the Indian original is uncertain, but it was probably Brahmasphuta Siddhānta of Brahmagupta written in 628 AD (No170; Bo250). In Arabic, it was called the Zīj al-Sindhind. Ptolemy's *Tetrabiblos* was translated from Greek circa 780.

The translation work under the caliphs, particularly by Caliphs Hārūn al-Rashīd (r. 786–809 [La66]) and al-Ma'mūn (r. 813–833 [La66]) (No179). Caliph al-Ma'mūn founded an academy with an observatory, the *Bayt al-Hikma* (House of Wisdom) to promote the translation of scientific and philosophical work into Arabic principally from Greek (La83; No179). Al-Ma'mūn is believed to have sought out Greek manuscripts from the the Byzantine empire (No179). Teams of translators were used and they compared their efforts to earlier translations made from Syriac versions of Greek texts.

Many important translations were done in Baghdad by a translation school (which replaced the House of Wisdom) guided by Hunain Ibn Ishāq (d. 873 [Sa248]) (La94). Hunain made a good translation of the *Almagest* (Pe154) which was later revised by the astronomer Thābit ibn Qurra (c. 826–901) (Pe161). Hunain, who was most interested in medicine, devoted much of his career to translating the Greek physicians, Hippocrates, Galen, and Dioscorides (Sa249; Gi44). As was not unusual at this early time in Islamic history, neither Hunain nor Thābit were Muslims: Hunain was a Nestorian Christian and Thābit, remarkably, was a pagan from Harran, a city that was the center of an astral religion (Gi44).

The translation movement as a whole was quite amazing. Much of the best of Greek science was now available in the new international language of Arabic. The reasons for this particular achievement cannot be easily separated from the reasons that the Caliphal court supported a wide range of activity in the arts and sciences. Lapidus offers the following analysis of reasons for Caliphal patronage:

> Futhermore all the court literatures served to propagate a pre-Islamic concept of the ruler and the empire. Interest in the secular aspects of Arabic literature, Persian adab, and Hellenistic philosophies and sciences signified the appropriation of a cultural heritage which could be used to legitimize Caliphal rule. They provided, in the Arabic case, an ethnic concept of political leadership; in the Persian case, a continuation of the heritage of ancient Middle Eastern kings; and in the Hellenistic case, a concept of the structure of the universe itself, in philosophic and scientific form, as the ultimate justification for imperial rule. The patronage of these several literatures implied ultimately that the Caliph, though a Muslim ruler, was legitimized in non-Islamic cultural terms going back to the heritage of the ancient Middle East (La97).

One could make a somewhat, but not altogether, different statement: the Caliphal court patronized the arts and sciences because they thought it their duty to enrich the culture of the Arabic empire and simultaneously to prove their right to rule through their magnificence and achievements equalling or surpassing those of the past. Yet another and never to be underestimated reason is simply that the Caliphs and their courts were probably sometimes intensely and personally interested in the arts and science. The interplay of motivations was probably complex.

For the translation of scientific works, Saliba tentatively offers a particular additional explanation (Sa52). The Arab-speaking ruling class desired to end the exclusion of their members from civil service positions requiring particular expertise. Thus the translation of advanced works was encouraged to redress this exclusion. Saliba believes, however, that the translation of the scientific works still requires a full explanation.

Naturally, for science and natural philosophy to be pursued there had to be individuals who wished to do so. The Caliphal courts would probably not have patronized science if they could not find or if they were not sought out by natural philosophers. Such individuals certainly existed in the Indian, Persian, and Greek worlds all brought into contact or, indeed partially encorporated, in the Arabic empire. One can mention in particular the school of Jundīshāpūr, a Sasanian Persian royal city (La93–94; No179). This school of Greek of philosophy was descendant of pagan Greek schools some of whose members who fled persecution in the Christianized Byzantine empire. The Nestorian Christian church of the Middle East, however, supported the philosophers and established them or their followers eventually in the 6th century in Jundīshāpūr. In the Islamic period the Jundīshāpūr school, most of whose members were Nestorian Christians, was transferred to Baghdad (La94).

The support of the Nestorian Christian church however helpful seemed insufficient for more than very sporadic science or dormant science since nothing much is said of the achievements of science in the immediate pre-Islamic Middel East. The patronage of the Caliphal court seemed necessary for the resurgence of scientific activity. Moreover, the need for princely patronage (first from Caliphs and later from other potentates) seems to have been a constant of Islamic science (Co386). Certainly, the four greatest centers of astronomical achievement, 9th-10th century Baghdad, 13th century Marāgha, 13th century Castile, and 15th century Samarkand were all made possible by great patrons: the caliphs, Hulagu Khan, King Alfonso, and Uleg Beg.

A possible reason for the decline of Islamic science after circa 1500 is that purely by the chance, there simply ceased to be even a handful of sultans and princes with scientific interest. And those that were such as Jai Singh in 18th century India (No201) may have been so unlucky as to discover no persons of genius to fulfill their ambitions. Given that there were no other institutions in Islamic society besides the court to pay, provide instruments for, to encourage, to pet and stroke scientists, science became dormant.

This explanation is certainly in accord with the Ben-David model: science is weak in traditional societies because support is weak and uncertain. However, one wonders why no other institutions besides the courts were

The 9th and 10th centuries in the Islamic world have been called the Golden Age of Islamic science when clusters of creative scientist were acquiring the Greek and Indian heritages and building on them.

This Greek heritage was enriched by the achievements of Indian mathematics and Hindu-Arabic numerals. The Islamic scientists such

The period 800–1000 has been called the Golden Age of Islamic science because of the clusters of eminent scientist who flourished then; the later age would then be called one of sporadic science (Sayili 1960; see also Co398). This neat division has been disputed (Sa8, 31). Probably the periodization needs more study. The last great site of Islamic science was probably the observatory near Samarkand of Uleg Beg (1394–1449) (No200; Pe398; La281). Later observatories were founded in Istanbul (1574) and in India in the 18th century by Jai Singh (No200–201), but they did not achieve much.

The essential support of science always seems to have been court patronage (Co386;

As in the case of Greek science, the decline of Islamic astronomy cannot isolated from the decline of Islamic science as a whole.

The Islamic madrasas were institutionalized schools for higher learning. Unlike the European universities (which may have been partially modeled on madrasas), madrasas generally did not teach the *awâil* sciences, but concentrated on Islamic law and theology (Co386, 398–399; La324). There were exceptions: in the early Ottoman empire, the lowest-level madrasas included logic, geometry, and astronomy in their curricula; the higher level madrasas did not however. By the 1540s, however, *awâil* sciences were being excluded and the narrow focus on Islamic studies (Quran, hadith [authoritative reports of early practice], and Shari'a [Islamic law]) was gaining ground (La327).

The Istanbul observatory attached to the Sulaymaniya madrasa was founded in 1574–1575, but demolished already in 1580 (No201; La327). This serves as example of the continuing astronomical tradition, but also of its weakness in the intellectual world of post-Medieval Islam.

In the 18th through 20th centuries there was a revival of interest in Islamic society in science and technology (e.g., La342, 620, 728, 814). The revival, however, was not primarily an indigenous phenomena, but rather a reaction to the challenge of European society. Reasonably enough, the Islamic science of the Medieval Ages could be taken as a precedent: an ancient tradition of Islam that should never have been cut from the vine (e.g., La620).

6. The Copernican Revolution

The Copernican revolution is an expression with various usuages. Here we take it to mean the transformation in astronomy and cosmology which took place in the period circa 1500–1700. This transformation replaced the Aristotelian-Ptolemaic universe with the Copernican-Keplerian-Newtonian universe The former was geocentric, finite, and based on epicycle-deferent models with the heavenly bodies being carried around by Aristotelian ethereal celestial spheres. The physics of the of Aristotelian-Ptolemaic universe was Aristotelian and was divided into two distinct realms: the heavens and the Earth (or more exactly the sublunary sphere). The later made the Earth a planet and had the planets going the Sun in nearly elliptical orbits. The Copernican-Keplerian-Newtonian universe was certainly large, possibly infinite, had the stars at immense distances, and made the Sun a star. The physics of this new universe was Newtonian: the same for the heavens and Earth.

The Copernican revolution is part of a larger transformation: the Scientific Revolution. Its role in the Scientific Revolution is plausibly the largest and showiest one, but it is not an indispensible role. Since the Copernican revolution makes most sense in the context of the Scientific Revoltion, we will discuss the later in general at some length in § 6.1.

6.1 The Scientific Revolution

We discussed the Scientific Revolution briefly in § 1.2. Here we expand on the theme in four subsections. The first discusses in general the dynamism of European society emerging out of the medieval period. The second discusses this dynamism with reference to particular innovations up till the eve of the Scientific Revolution. In the third subsection, we discuss the intellectual effects of this dynamism in the period of the Scientific Revolution. The fourth subsection takes up the nature of the Scientific Revolution. It should be emphasized that my discussions are very simplified and are largely based on the ideas of and cited by Cohen in his book *The Scientific Revolution: A Historiographical Inquiry*. The particular synthesis of the elements is my own, however, and there are many important and interesting elements not mentioned

6.1.1 The Dynamism of Europe

In discussing the background of the Scientific Revolution, we are really discussing its external causes: i.e., those causes from outside of the purely scientific realm. Such a discussion is clearly uncertain ground of historical explanation. One thing is clear, the Scientific Revolution arose out of a complex of causes.

The factor, the precondition, I believe of overwhelming importance to the initiation of the Scientific Revolution is dynamism of Europe (here meaning mostly western Europe) developing in the centuries preceding the Scientific Revolution and continuing through it (and ever after as well). This dynamism consists in the rapid technological, economic, and scholarly developments in Europe emerging from the Middle Ages.

One of the obvious things about human history and much of prehistory the cumulative development of technology. Technological progress was usually so slow as to excite no comment. In a particular craft, a person might recognize the innovations seen in their lifetime, but an awareness of general innovation was not so obvious. But retrospectively, we can recognize it occurring. There were certainly periods of technological decline. The large civil engineering technologies of the Roman empire disappeared at least in western Europe with the decline and fall of the Western Roman empire. However, such declines are often local in scope; summing over all societies at any one time, significant net declines seem to be rare.

The progressive (if unsteady) technological development is not mysterious. Given stable social conditions (e.g., no devastating barbarian invasions), useful technological skills tend to be maintained and useful innovations retained. The transmission of skills, until modern times, was largely by apprenticeship rather than in special schools or from books.

Circa 1000, European society, just emerging from the Dark Ages was not at the technological forefront of human societies. Islamic, Indian, and Chinese societies were much in the lead. However, out of the Dark Ages Europe emerged with robust agricultural technology that could support a large, complex society (Gies44, 109). The evidence of what in fact happened shows that this early medieval society was eager to learn and utilize technologies from abroad (Gies82), and then to improve on them.

It often cited that Europe's pluralism (e.g., La269) was a great advantage in rapid technological development. In a heavily bureaucratic society, innovation can be ordered or forbidden top-down. If such a society turns conservative and scorns outside inventions (as China did [Car32]), then technological development can slow. Medieval Europe on the other hand exhibited an often cited pluralism (e.g., La269). It was divided into a number of competing states and within states feudal divisions worked against centralization. But aside from state, society was divided into potentially innovative components that could act rather independently: e.g., feudal lords, the Church hierarchy, the monastic orders, a somewhat free peasantry, and some very important self-governing cities and town. If a technological innovation was missed in one place, it could be picked up in another, and then imitation and competition might cause its general adoption.

The European society was also safe for innovation. In any society riven by wars and massive raids any technological innovation demanding capital expense is unlikely: the warriors will just take what they can (including people for slaves) and destroy at least some of what they cannot. There were continual feudal wars, of course, but these do not seem to have impaired technological and economic development. European feudal wars seem to have been rather limited in scope. Various princes and kings exchanged territories in a hostile fashion, but they did not generally aim at annihilating rivals, who were often cousins. The elite arm of their armies were knights and their commanders were barons. These were not landless warriors, but persons of substance who had a vested interest in not destabilizing the overall system by fighting for crushing victories and who did not want overmighty rulers.

Fortunately, Europe suffered no devastating invasions from without after the 9th and 10th century struggles with the Vikings and Magyars (Co407; Gies42). At the European southern periphery, European forces stemmed the tide of Muslim conquest in Spain and drove the Muslim regimes out in the course of the Middle Ages (La379, 383–384). (The latter was human tragedy as we now understand and arguably also a tragedy for Spanish culture: heralding the doom of its pluralism, openness, and toleration.) A bad moment for Europe came with the Mongol invasion of 1241 under Batu Khan (another grandson of Genghis Khan and the founder of the Khanate of the Golden Horde) (Ru251–252, 294). Batu's forces soundly defeated Polish and Hungarian armies. It seems that Batu may have wished to stay in Europe, but the death of the Supreme Khan in the Far East caused him to retreat with his army: he wished to play a forceful role in the succession decision. It is possible that the Mongols would have wreaked devastation on Europe and perhaps have established a khanate there that would have changed decisively Europe's development. On the other hand perhaps Mongols were overextending themselves; they never came back.

A constrasting case is presented by the Middle East. In the late 9th and early 10th centuries, there was almost constant warfare and (not unrelatedly) heavy fiscal exploitation by governments (mostly ephemeral); this led to an economic regression (La133–136). The traditional military organization that developed in the later Caliphate and that was retained for centuries was based on slave armies (La127, 148). Such armies had no special loyalty to society, only to their particular ruler or to their own self-interest. With such a military organization wars would tend to be more extreme than in Europe and military requisitioning in war and peace more exploitive. After 1000, there were two great waves of invasions: the Turks from inner Asia beginning in the 11th century and the Mongols in the 13th. The Turks, of course, quickly converted to Islam, but their immigration was at least initially destructive and it continued for centuries (La142, 306). The Mongol destructions we have already remarked on in § 5.5.1. It has been suggested by J.J. Saunders (see Co408–409) as possible that, besides the material destruction hindering all development, the invasions changed the temper of Islamic society. The confident outlook on other cultures based on the physical strength of early Caliphal period was transmuted into a conservative inward-looking view: the continuance of Islam needed a tight hand on the past, not an imaginative reach for new enterprises of the mind. Whether this suggestion has some truth or not, Europe was saved from being another possible case inward-looking conservativism.

Medieval Europe did suffer one major disaster: the Black Death. Beginning in 1348, the Black Death or plague killed perhaps a third of Europe's population by 1400????. Some historians perceive a slowing in technological innovation in the 14th century and suggest the plague as a possible cause (Car49). It is easy to imagine that the uncertainties of life at that time were increased and that markets and workforces were changing unpredictably; so it is possible that the plague curbed Europe's progress. However, the plague did not destroy material wealth and though it affected all limbs and organs of society it did not excise any of them. Europe's population afterward was still large and in fact suffered less population pressure?????. Although personally horrific, the plague does not seem to have tremedously damaged European society.

On the intellectual front, medieval Europe was also progressing. Circa 1000, the book learning of Europe was meager compared to antiquity and to Islamic world. However, from the Islamic world and from Greek manuscripts, the ancient heritage plus some additions of Islamic society and farther east were obtained and translated to Latin, the international and scholarly language of Europe. The 12th century saw the beginning of translations from Arabic and Greek (Pe167ff). The Arabic translations were mostly done in Spain and Sicily, and unfortunately declined after the 12th century.

The most important of these recoveries for the medieval scholars were the works of Aristotle. To them Aristotle became the prime authority: he was 'the Philosopher' (Co158). It is interesting to note that Aristotelianism was a radical innovation in the universities (see below) in the 13th century: the dogma of his authority developed quickly however (Pe168ff). Some important works of Archimedes were translated from the Greek in the 13th century by Willem van Moerbeke (c. 1215–1286 [Pe401]), but they failed to stimulate an interest in Archimedian (i.e., a mathematical) physics at that time (Co275). Perhaps the most important work of physics from the Islamic world was the *Optics* Ibn al-Haytham (c. 965–1039 [Pe349]) translated in the 12th century (Pe163ff, 167). New numerical techniques and Hindu-Arabic numerals (see § 5.2) also arrived Europe from the Islamic world. The latter came gradually into general use over the course of the 13th and 14th centuries.

The scholarly component of medieval society was in the Church. The Church provided the scholars and students in the main for the universities, a European innovation (Car44). The universities were not primarily aimed at innovation themselves though, but at education. Fortunately, the undergraduate education was in principle a broad one: the seven liberals arts constituted out of the trivium (grammar, rhetoric, and logic) and the quadrivium (music, geometry, arithmetic, and astronomy). Beyond the level of arts faculty were the three professional faculties: divinity, law, and medicine. The core philosophy of the universities was Aristotelianism from the 13th century onward as we noted above. Despite being not aimed at innovation, some advances in philosophy and natural philosophy did occur in the universities: e.g., at Oxford (Pe193) and the University of Paris (Pe196). The university education did provide a springboard for many later developments and most of the leaders of the Scientific Revolution were university trained although they rebelled against the Aristotelian orthodoxy.

It is probably very important for the Scientific Revolution that the sciences were an institutionalized part of the university curriculum. This gave the sciences a respectable place in society, independent of the favor of princes and with the explicit approval of the religious establishment. The contrasting situation in the Islamic world, we discussed in § ???. Despite the famous conflicts of religion and science in European history, it is essentially true that western Christianity has generally approved of science since 1000 at least. Without this approval science may have found itself marginalized and suspect. Whether such an unhappy state could have ruled out a Scientific Revolution is impossible to say (see discussion in Co384–417), but it certainly would have militated against a Scientific Revolution.

6.1.2 Particular Expressions of the European Dynamism

The technological state of Europe circa 1500 was as advanced as the world had ever seen. Only China long ahead in technology was rival. The technology of the Greco-Roman civilization had been surpassed in all but a few special cases. In science, the distinction from Greco-Roman times was not so great.

The technological transformation of Europe since Greco-Roman times was immense and multifarious (Gies15). For example, the Greco-Romans had never exploited water power and horse power fully (Gies32ff). Waterwheels were known, but their usage scarce (Gies35). There is only one possible reference from late Antiquity to using a waterwheel for any other function than grinding grain (Gies35). But by 1500, the waterwheel powered metallurgy, sawmills, textile production; it helped to make olive oil, beer, mustard, paper; it ventilated and pumped out mines (Gies265). In Roman times, the horse harnesses in use tended to strangle horses and horses worked without horseshoes which are needed to protect hooves especially in moist climates (Gies46; Lynn 1966, p. 308–309). Horses, can work longer, move faster than oxen; with the padded horse collar and horseshoe (both arriving in Europe before 1000), horses became the other main power source of Europe both at the plow and in transport (Gies45–46, 149).

Two other innovations of immense importance were printing, and gunpowder and guns. Movable type printing probably began in Europe with Johann Gutenberg in Mainz circa 1450 (Gies241ff). The effects were quickly seen: by 1500, there were $\sim 40,000$ editions and $\sim$ 15–20 million copies. In the first 50 years after the invention in 1450, more books were produced in Europe than in the previous 1000 years (Car55). Passing into circulation were the works of classical and medieval scholars, religious tracts, text books of all sorts (e.g., mathematics, engineering), accounts of the oceanic voyages, literature. The invention of the printing press was expanding the intellectual empire of every literate person. Intellectual chain reactions were set off impossible to simply describe.

Guns and gunpowder gave a destructive power not seen before. Cannons could knock down medieval castles and being expensive could in quantity only be afforded by the state in general (Gies249ff). The small, locally powerful medieval lord could not maintain military independence. The power of the state was enhanced. The knight with his individual prowess was replaced by mass armies with muskets and pistols who literally gave much more bang for the buck. The feudal age, probably already in decline, could not resist. Perhaps the actual waste of war did not increase so much by 1500 because of guns, but in later centuries with evermore gun development, it certainly did. And of course, gunpower gave the European conquistadors their most obvious and immediate tool. The age of *Bellifortis* (strong war) had begun.

There are other notable innovations. Lenses??? and spectacles first appeared in Italy before 1292 (Gies227). The earliest lenses were only convex and so only a boon for the the farsighted; concave lenses (good for myopia) came in the 16th??? century. Mechanical clocks were first known from circa 1280???. However, not until the 17th century were they accurate enough for important astronomical purposes. Indian-Arabic numerals became well known in Europe after 1200: there advantages both for business and science over Roman numerals gradually became evident (Gies225-227). Paper which originated in China, was being imported to Europe by 1000 and manufactured there by 1200????.

Many of the innovations cited above were not European in origin: they came from China, India, the Middle East, and elsewhere. The actual history of transmission is often obscure: the spread of paper from China is one of the most easily traced. Other innovations (e.g., spectacles) clearly originated in Europe. Of course, independent innovations in the east and Europe also certainly occurred. It is difficult to decide often between transmission and independent invention given the patchy history of pre-modern technology.

If technology was far in advance of the Greco-Roman world, science or natural philosophy (i.e., an understanding of the world into terms of elementary principles or laws) was not so. However, as we discussed in § 1.1, much of the best of ancient science had been recovered, and some of Islamic science had been obtained. It is a pity that the work of the Marāgha astronomers and Uleg Beg's observatory was not translated.

In the religious sphere a mammoth change occurred right at the start of the Scientific Revolution: the Protestant Reformation. The roots of this change probably lie in ending of the clerical monopoly in learning (which was aided by the growing prosperity and sophistication of the lay class) and the intellectual endeavors of the humanist movement aided by the printing press. The humanist movement especially in the printed and best-selling writings of Desiderius Erasmus (1466–1536 [Ba407]), had challenged the medieval Biblical interpretation. The changed perspective on the Bible led many to conclude that the Church of Rome had no real authority and in fact was greatly in error. Although many other complex causes can be cited this seems the most essential cause of the Reformation which started with Martin Luther's 95 theses in 1517. The cracking of the unity of western Christendom had fateful consequences, but its effects on the Scientific Revolution are hard to untangle. Leaders of the Scientific Revolution were found on both sides of the Catholic-Protestant divide.

This is not the place to expand on the artistic achievements of the Middle Ages and the Renaissance. However, they were great. The Middle Ages provided most notably cathedrals. In the Renaissance period the creations of Leonardo da Vinci (1452–1519 [Ba1358]), Michelangelo (1475–1564 [Ba768]), Raphael (1483–1520 [Ba1003]), Albrecht Dürer (1471–1528 [Ba375]) and many others hardly need to be commented on. Their contemporaries recognized them as great masters and knew them to be at least on par with the artists of antiquity.

Then there were the oceanic voyages of Columbus (1446?–1506 [Ba239]), Vasco da Gama (1469?–1524 [Ba498]), and others. These voyages verified the sphericity of the world, brought the continents into contact, and started trains of developments and disasters (intentional and unintentional). They dramatically demonstrated that the ancients had not known everything: the *Geography* of Ptolemy was wrong (which was hardly surprising) and that the equatorial zone of the Earth was not impassible due to extreme head. The voyages also promised economic gain which did not fail to materialize. The oceanic voyages were made possible in the 15th century by the earlier developments in maritime technology. The full-rigged ship made wind power more tractable than ever before (Gies227ff). Compasses, charts, and chartmaking improved navigation: you knew where you were going (if it was not truly the first voyage) and sometimes you even knew where you were (Gies278). Latitude from the declination of the Sun or stars and time of year was easily determined. But longitude at sea could not be determined accurately before the invention of chronometers in the 18th century (Gies279).

6.1.3 The Intellectual Effects of the European Dynamism

The 14th and 15th centuries, besides being a time of change were also a time of perceived changed. People knew things unheard of in their grandparents days. It was also clear that the ancients had been surpassed in at least in technology and geography. The European dynamism of was unquestionably stimulating. Widespread change and improvement was possible; why could it not be willed, created, made a permanent feature of life? To quote Francis Bacon (1561–1626 [Ba91]), the prophet of progress in science and in technology through science, had this to say:

> ... by the distant voyages and travels that have become frequent in our times many things in nature have been laid open and discovered which may let in new life upon philosophy. And surely it would be disgraceful if while regions of the material globe—that is, of the Earth, of the sea, and of the stars—have been in our times laid widely open and revealed, the intellectual globe should remain shut up within the narrow limits of old discoveries (quoted from Car77).

Bacon was writing after 1605 (Car76), but views he expresses must have been held by others much earlier.

The ancient philosophers did not know everything; in fact they were known to have been wrong on points. The seemingly static philosophy of Aristotelianism need not be the last word. The Scientific Revolution provided the new word.

6.1.4 The Nature of the Scientific Revolution

The expression, Scientific Revolution, in its modern historiographical usage goes back to the

1930's and the historian Alexandre Koyré (Co21). Its exact meaning and ideas as to how it came about are the topics of Cohen's book (Co).

An "average view" is that the Scientific Revolution occurred over the time span circa 1500–1700. There are an abundance of ideas about exactly what constitutes the Scientific Revolution and explanations for it. Koyré and others emphasized the intellectual transformation, in Koyré's words: "from the world of the 'more-or-less' to the universe of precision" (quoted from Co86). In other words, nature came to seen as being governed by mathematical laws generally and these laws were discoverable.

Naturally, a mathematical nature was not altogether new. Astronomy had been mathematical in a crude sense even in prehistory (No3) and had become much more so in the times of the ancient Mesopotamians and Greeks. The mathematics of the vibrating string had been known since the Pythagoreans (i.e., the 5th century) (Pe17). Aristotelian physics (c. 4th century BC), although largely qualitative, did include the theorem of the parallelogram of velocities (Pe106). Archimedes (287?–212 BC [Pe308]) had mathematized statics (Pe90ff) and hydrostatics (Pe99ff). But despite these important examples, the nature in general was not perceived as being strictly mathematical. Aristotelian physics, widely accepted in the Greco-Roman, Islamic, medieval European, and Renaissance times, was essentially qualitative, especially notably in dynamics (Pe105ff).

The new thrust in the 'mathematization of nature' began with the heliocentrism of Copernicus. The essential boon of heliocentrism in Copernicus' view, was not that it was provable, for he could not prove it, but that it implied "the structure of the universe and the true symmetry of its parts" (Ro4; see also Ku171, 174–175). Copernicus thus propounded symmetry (which is a mathematical principle among other things) as a guide in natural philosophy. Moreover, his theory was radically inconsistent with Aristotelian physics: Copernicanism implied that the physics of the heavens and Earth were one. Kepler and Galileo would go much further pursuing mathematical symmetries and a unified mathematical physics. Newton would cap the process by providing a mathematical physics from which heavenly and Earthly dynamics could be derived quantitatively. Thus, qualitative understanding was replaced by a much more exact quantitative understanding. The tremendous success of the Copernican revolution in providing a correcter view of nature, was the main cause in making mathematical understanding the ideal that has been pursued ever since.

The mathematization of nature was not the only essential thrust of the Scientific Revolution. There was also development of the doctrine of detailed observation and experimentation: the 'new empiricism.' There was, of course, an 'old empiricism.' Ancient mathematical astronomy had been based on detailed observation. Aristotle had considered experience the foundation of knowledge and had emphasized careful observation (Pe30–31). In the Islamic world there had been a re-emphasis on empiricism as evidenced by some improvements in observational astronomy over the ancients (e.g., No187, 200). But despite Aristotle's empiricism, his ancient and medieval followers usually saw no reason to go beyond casual qualitative observations; for this reason, they could only make limited progress beyond Aristotle, and, in fact, progress was not a goal for many of them. As for the Islamic empiricism, its bolt had been shot with the decline of Islamic science (see § 5.7).

The new empiricism is not, of course, unrelated to mathematization. The mathematization of nature to proceed as it did needed to be guided and checked by a mathematical and precise observation of nature which whenever possible (and it is generally not possible in astronomy) was to be extended to new realms by experimentation. Kepler's theoretical work was based on the unprecendented accuracy of Tycho Brahe's astronomical observations (No302). The role of experimentation in Galileo's science has been much debated (Co108ff). But it clearly was an essential element, equal to that of theory as the researches of Drake (e.g., Dr: see esp. Dr408–410) and others have shown (Co108ff, 132).

The new empiricism was in fact broader in scope than the mathematization of nature since it could be applied to the sciences not easily amenable to mathematization: the life sciences, and electrical and magnetical physics (which were not mathematical in the 16th–17th centuries).

The achievements of the early modern scientists were so impressive that their method, essentially the modern scientific method became the established procedure for science. The atitude that found science intellectually rewarding also seems to have spread and increased.

Additionally, the notion that science would lead to a physical mastery of nature became widespread through the writings of Francis Bacon (1561–1626 [Ba91]), René Descartes (1596–1650 [Ba327]), and others (Co191ff). This idea, although not wholely absent in ancient times, became much, much more current. This idea of applied science undoubtably increased the support for science somewhat. But this assertion must be tempered by the fact that science-based technology only slowly developed; the material rewards of science did not flow like liquid gold in the 16th–17th centuries, and did not really do so until the 19th (Co195). Some inventions did come from the hands of scientists, notably scientific instruments (Co195), but by and large technologists were still solving their own problems in the period of the Scientific Revolution. Only gradually, in the 17th-18th centuries did individuals appear who provided the link between science and technology (Co194).

The intellectual power and achievements of early modern science bolstered somewhat by the foreseen, but not much realized, material benefits of science are probably the main reasons for the strong grip that science acquired in Europe. The mold of science in traditional societies was broken; science became continuously progressive. It is the transformation to continuously progressive science and the emergence of the modern scientific method which forms the essence of the Scientific Revolution. The fact that science and technology (in a complex symbiosis) have transformed the intellectual and material world again and again since the time of 16th century and give no indication of halting, makes the Scientific Revolution one of the outstanding events of history.

6.2 Copernicus

Nicolas Copernicus (1473–1543 [Pe324]) was born in Torun (in German Thorn) on the Vistula River in Poland. He was a child of the Renaissance. His dates make him a contemporary of Christopher Columbus (1446?–1506 [Ba239]), Leonardo da Vinci (1452–1519 [Ba1358]), Montezuma (1477?–1520 [Ba789]), Martin Luther (1483–1546 [Ba727]), and Lucrezia Borgia (1480–1519 [Ba139]).

There has always been some controversy about Copernicus' nationality (Ko132). He was born and lived most of his life in what is now modern Poland and was then part of the Polish kingdom. In some periods between those times some of the territories he lived in were parts of Prussia and Germany. He considered himself a subject of the Polish king (Ko132???), but his native language was German and he enrolled in the German Nation (a student society) when studying at the University of Bologna (Ros127). It is probably wisest to say he was both Polish and German.

Copernicus' name is also a bit debatable. In his times, the spellings of names were not often fixed. In documents his name most often appears in German form as Koppernigk (Ko596). Copernicus, himself, signed his name variously, but seems to have used the latinized Copernicus mostly in later life (Ko569). Thus, it is fair for history to adopt Copernicus.

Copernicus' father, a wealthy merchant, died in 1483 and Copernicus was subsequently protected and patronized by his uncle Lucas Watzenrode who became bishop of Ermland (Varmia) in 1494 (Pe324; Ko128–129; Ada107). Copernicus studied at the University of Cracow in the years 1491–1495????, but left without taking a degree (Ko130; Ros 127). In the years 1496–1503, Copernicus was mostly in Italy studying at the Universities of Bologna and Padua (Ko131, 137; Pa189; Ada101). He studied a wide range of subjects including mathematics, astronomy, Greek, philosophy, medicine, and law (Ko131). Medicine was particularly important for Copernicus and he was noted as a physician (Ko134, 137), but he took no degree in that subject. His only degree was a doctorate in canon law granted May 31 1503 by the University of Ferrara where in fact he had never studied (Ko131; Ros179). He took his examinations and degree there because the examinations were easier, and the cost of the degree itself and ancillary fees (for celebrations) were less???? (Ko132; Ros179; Ada97).

At the University of Bologna, Copernicus studied with one of the leading astronomers of the day, Domenico Maria Novara (1454–1504) (Ros129). Novara, professor at Bologna from 1483 until his death, was an original thinker and was bold enough to challenge ancient authority. Unfortunately for him, one of his most notable discoveries was quite wrong; he thought he had discovered that the north celestial pole had shifted and the latitude of European positions was further north by $\sim 1^\circ$ than (Ros132–133). In regard to the planetary system, however, Novara remained a traditional geocentrist (Ros 130).

In 1497, Copernicus' uncle appointed Copernicus a canon of the Cathedral of Frauenburg in Ermland (in Polish, Frombork in Varmia) (Ko131). This appointment (literal nepotism) was for life and thus gave Copernicus security and a good income (Ko143–144). His duties as a canon were essentially to help administer the estates of the cathedral; he was also expected to attend morning and evening services (Ko144). Through most of history, canons have usually been priests???? and some of a canon's duties can only be performed by a priest (e.g., celebrating mass). In Frauenburg, however, a substitute could be employed to carry out the priestly duties. In Copernicus' time many of the Frauenburg canons were not priests (Ko144; Ada102), despite some efforts to force ordination (Ada102).

Copernicus never became a priest it seems (Ros47ff). Copernicus is never referred to as priest by any of his contemporaries or anyone for 70 years after his death (Ros47). In at least one document where it might expected that he be referred to as a priest if had been one, he not so referred to (Ros50). However, the story that he had been a priest has had a long life. The first ordination of Copernicus was by Galileo in his *Letter to the Grand Duchess Christina* in which Galileo makes several historical errors (Ros193–204). These errors can be attributed (almost entirely anyway) to the level of historical scholarship of Galileo's day and Galileo's willingness to assume that Copernicus was as Catholic as possible, to help make Copernicanism as acceptably Catholic as possible (Ros204). Galileo probably made the assumption that Copernicus had been a priest because Copernicus had been a canon. And seventy years after Copernicus' death and out of Poland, there were probably few who could have certainly denied the assertion that Copernicus was priest. Galileo's erroneous statements became an infection in the historical literature that has only recently been cleared up (Ros193–204). It is, of course, barely possible that Copernicus was a priest because there is no contemporary source absolutely asserting he was not.

Copernicus did not reside in Frauenburg immediately upon his canonship. He was mostly studying in Italy until 1503. In the period 1503–1512, he was with his uncle, the bishop of Ermland, living in Heilsberg Castle (Lidzbark Castle in Polish [Ada109]), the episcopal residence (Ko137). He was officially a physician to the bishop, but also acted as a an assistant to the bishop in diplomatic and other affairs (Ko138; Ada114). In 1512, Bishop Watzenrode died and Copernicus took up residence in Fraunberg which is located on the Frisches Haff (a large bay on the Baltic Sea) (Ko143). In the dedication to Pope Paul III (r. 1534–1550) of his book *De Revolutionibus Orbium Coelestium* (*On the Revolutions of the Celestial Spheres*), which will be abbreviated as *Revolutions,* Copernicus refers to Frauenberg as "this very remote corner of the earth" (Ro5). The remark was probably intended lightly, but it is true that Copernicus was physically removed from the great centers of learning and affairs. Copernicus lived in Frauenburg, performing his duties (mostly administrative) as a canon for the rest of his life (Ko143).

6.2.1 The Commentariolus

When Copernicus became converted to heliocentrism is unknown. His first work expounding heliocentrism was known to have existed and been distributed as of May 1 1514 (Ros71). This work is only a manuscript and does not give Copernicus' name (Ros71) or have a title (Ros73). An abbreviation of the title given to it by Tycho Brahe, the *Commentariolus* (little commentary) has become its common designation (Ros73).

In the *Commentariolus* there are 7 postulates which can be paraphrased (following Ros74–75 and Ko148–149) as follows:

1. The universe has no single center.
2. The Earth is not at the center of the universe.
3. The center of the universe is close to the Sun.
4. The Sun-Earth distance is negligible compared to the distances to the stars.
5. The apparent daily westward rotation of the Celestial sphere is do to the daily eastward rotation of the Earth on its axis.

6. The apparent yearly revolution of the Sun is caused by the Earth's revolution about the Sun.

7. The planets also revolve around the Sun and their retrograde motions are caused by the relative motions of the Earth and the planets.

These postulates are not stated as compactly as they would be by a modern. For example, the remark about retrograde motions is really a deduction. The postulates, however, give the essence of Copernicanism.

After the *Commentariolus*, Copernicus does not seem to have broadcast his ideas in anyway until 1540 (Ko166). Nonetheless seems that based on the *Commentariolus* and reports of it, Copernicus' ideas became somewhat broadly, if not well, known. The Copernican idea was explained in lecture in the Vatican gardens to Pope Clement VII (r. 1523–1534) in 1533 (Ros153; Ko155–156). Clement reacted favorably, but there is no evidence that Copernicus ever knew of the lecture or Clement's reaction. Nicolas Schönberg, the Cardinal of Capua and advisor to Popes Leo X (r. 1513–1522), Clement VII, and Paul III, having heard reports of Copernicus' ideas, urged him to publish in 1536 (Ko154–155; RoXVII). Less happy notice was also taken. In 1539, Martin Luther, after mention of an unnamed astronomer and his theories, is quoted as having said:

> But that is how things go nowadays. Anyone who wants to be clever must not let himself like (sic) what others do. He must produce his own product, as this fool (*Narr* in the original German) does, who wishes to turn the whole of astronomy upside down. But I believe in Holy Scripture, since Joshua ordered the sun, not the earth, to stand still (adapted from Ros197–198; see also Ku191 and Ko156–157, 572).

In justice to Luther, his statement was apparently an offhand remark, not an authoritative one. Copernicus was also apparently mocked, presumably for his theories, in carnival farce in the Prussian city of Elbing in 1531 or so (Ko156).

6.2.2 The Delay and Publication

One wonders why Copernicus did not speedily follow up the the *Commentariolus* with a published treatise. Probably one reason is that his ideas were not yet ripe. The *Revolutions*, Copernicus' only published book on astronomy and his theory is a large work filled with detailed proofs and calculations. It took years to prepare this work and Copernicus, knowing that he was presenting radical ideas, may not have wished to go public with anything less than the most solid, definitive presentation he could give. Rosen contends that Copernicus was working on the *Revolutions* from no later than 1515 (Ros109), that he worked on it for years (parts being written before and after about 1524 [Ros114]), and that the final version was probably finished in 1541 (Ros116).

Related to the first reason was probably Copernicus' rather retiring nature.??? Copernicanism was a direct challenge to the Aristotelian-Ptolemaic orthodoxy of the time. Copernicus could well imagine that he would be howled down (as indeed he was after his death) by the orthodox natural philosophers. Staying out of wrangles may have been a considerable motivation to keep silent until late in life.

A third reason, that may have been the most important one, was fear of condemnation for heresy. The Bible contains several passages that imply a moving Sun and resting Earth. The most famous and clear of these is the Joshua passage (Joshua 10:12–14) that Luther alluded to:

> Then spake Joshua to the Lord in the day when the Lord delivered up the Amorites before the children of Israel, and he said in the sight of Israel, Sun, stand thou still upon Gibeon; and thou, Moon, in the valley of Ajalon.
>
> And the sun stood still, and the moon stayed, until the people had avenged themselves upon their enemies. Is not this written in the book of Jasher? So the sun stood still in the midst of heaven, and hasted not to go down about a whole day.
>
> And there was no day like that before it or after it, that the Lord hearkened unto the

voice of a man.

This passage states that the Sun moves. Other passages imply a stationary Earth (e.g., Psalm 104 and Job 38). A modern exegete would have no trouble reconciling such passages on the grounds that all motion, at least kinematically, is relative. However, Copernicus' contemporaries and Copernicus himself thought in Aristotelian terms in which motion and rest are qualitatively distinct states; thus the modern escape was not available.

The reconciliation Scripture and Copernicanism promoted by the later Copernicans, Galileo and Kepler, is summarized in the remark attributed to Cardinal Baronio by Galileo: "the Holy Spirit is to teach us how one goes to heaven not how heaven goes" (Fa184; Ca299–300). The point is that the Bible is not a treatise on astronomy and speaks in the direct language of ordinary life; the Bible does not err, but it does not fully inform about subjects irrelevant to its purposes. Copernicus would have almost certainly have agreed; in his preface to the *Revolutions*, he speaks of those who will fault him by "badly distorting some passage of Scripture to their purpose" (Ro5).

The argument that the Bible is not speaking absolutely on astronomical points is easy for a modern Christian to accept. But it was not easy for many of Copernicus' contemporaries. The Bible (it seemed) and Aristotle agreed on the stationary Earth and the moving Sun. And Aristotle was the authority on natural philosophy. Thus the Bible was indeed speaking absolutely as verified by human reasoning at its best. Moreover, it was the opinion of some that the early Christian Fathers had agreed on immobility of the Earth and the motion of the Sun; after the Council of Trent, points of agreement among the Fathers was held to be authoritative by the Catholic Church (Fa231; Ko454). Moreover again there was common understanding (even apart from Aristotle): the Earth is at rest just as it seems. Thus, the Bible, Aristotle, Christian tradition, and common understanding concurred against Copernicanism.

Copernicus was fully aware that his ideas could be viewed as heretical. He had Cardinal Schönberg's letter expressing personal approval, but this was far from an official approval of the Catholic Church. Once Copernicanism was in print, Schönberg and other openminded high ecclesiastics may or may not have been able to protect it. The nature of Copernicus' time made that more difficult than at early times. The Protestant Reformation, started in 1517 with Luther's 95 theses, had caused a bitter divide in western Christianity with both Catholic and Protestant sides hurling anathemas at each other. In such an age radical ideas were bound to encounter bitter enemies. Possibly there would be public preachers on both sides of the Catholic-Protestant divide who would denounce Copernicanism. There could be an escalation of denunciation. Possibly, Copernicus would be called to Rome to answer questions from the Inquisition; he might have be forced to abjure his ideas and have been imprisoned. On the other hand perhaps nothing bad would have happened; even important points of astronomy do always cause public outcrys. The course of events was unpredictable. Copernicus probably felt it was better to delay.

In 1539, Georg Joachim Rheticus (1514–1576 [Ko157, 193]) visited Copernicus (Ko158). Inspired by only what was already public knowledge, Rheticus had converted enthusiastically to Copernicanism (Ko158). Rheticus and Copernicus' old friend Tiedemann Giese (1480–1550 [Ros58]), the bishop of Kulm (in Polish Chelmno) convinced Copernicus that his ideas should be presented openly. Rheticus prepared and printed in 1540 a small treatise summarizing the Copernican system, the *Narratio Prima*; Copernicus' was only referred to obliquely as Dr. Nicolas of Torun (Ko161, 166). The *Narratio Prima* was a trial balloon and a successful one; no lighning struck and there were requests from scholars for Copernicus to publish his own work (Ko166). Copernicus was at last ready; he was in his late 60s, and so perhaps he felt that time was running out for him and there was less to fear.

Copernicus' book, the *Revolutions,* was printed in Nuremburg (a Lutheran city) by the printer Johannes Petreius, a specialist in astronomical works (Ko168; RoXV). The work began in 1542 under Rheticus' direction, but he had to leave in November to take up a post of professor at Leipzig University (Ko169–170). Andreas Osiander, Protestant theologian favorably disposed to Copernicanism, took over the supervision of the printing (Ko170).

Still aware that the *Revolutions* could be attacked on Biblical grounds (as indeed it was to be), Copernicus attempted to protect it as far as possible. He prepended Cardinal Schönberg letter of

1536 urging publication (RoXVII; Ko154–155) and he dedicated the book to Pope Paul III. Tommaso Campanella (1568–1639 [Ros204]), a defender of Galileo, would assume, as perhaps Galileo also, that the dedication indicated that Paul III had approved the book; this assumption was based on Italian practice of Campanella's time (Ros204). In fact Paul III, did not approve the *Revolutions*.

In the dedicatory preface Copernicus warns against those ignorant of astronomy who will censure book on Scriptural grounds. He follows this with the statement that "Astronomy is written for astronomers" (Ros5). This remark seems arrogantly exclusive. But Copernicus' point was probably that the book should be read as astronomy, not as a Biblical exegesis. At the end of the preface suggests that his work will aid in the reform of the ecclesiastical calendar, a project supported by Leo X some years earlier.

Osiander on his own initiative added to the protection of the *Revolutions* by inserting his own anonymous preface before Schönberg's letter and Copernicus' preface (RoXVI). This preface ends with the paragraph:

> Therefore alongside the ancient hypotheses, which are no more probable, let us permit these new hypotheses also to become known, especially since they are admirable as well as simple and bring with them a huge treasure of very skillful observations. So far as hypotheses are concerned, let no one expect anything certain from astronomy, which cannot furnish it, lest he accept as the truth ideas conceived for another purpose, and depart this study a greater fool than when he entered it. Farewell (quoted from RoXVI; see also Ko573–574).

This preface reflects Osiander own skeptical view of the reality of astronomical models. In 1540, Osiander wrote to Copernicus as follows:

> For my part I have always felt about hypotheses that they are not articles of faith but bases of computation, so that even if they are false, it does not matter, provided that they exactly represent phenomena. ... it would therefore be a good thing if you could say something on this subject in your preface, for you would thus placate the Aristotelians and the theologians whose contradictions you fear (quoted from Ko171).

We do not know what Copernicus replied to this letter (Ko171) or if he approved any form of Osiander preface although he probably saw it in proof pages shortly before his death (Ko173–175).

It probable that Copernicus or any mathematical astronomer would have agreed with Osiander's remarks in respect to the details of epicycles and deferents of the astronomical models. Even to Ptolemy, if not before, it must have been clear that the models of the celestial motions with epicycles and deferents were not unique: some models were better than others, but no one had found a uniquely good system. But in regard to the main hypotheses, Ptolemy and Copernicus would not have shared Osiander's skeptism. They each believed that their main hypotheses were true or at least very defensible physically and philosophically (Pe87; Ro4–5). Osiander's skeptism was, however, widespread in his day. Strict Aristotelians for example had to adhere to it since they took Aristotle's model of the celestial system to be physically correct despite its many significant disagreements with observation (Pe69–70).

Osiander's preface was often taken to be from the hand of Copernicus despite the contradiction between it and the rest of the *Revolutions*. Some people such as the anti-Copernican Giovanni Maria Tolosani (1470?–1549 [Ros149]) recognized the preface as not by Copernicus (Ros157). Kepler in 1609 would expose Osiander's authorship (Ko172). The preface or perhaps just the assumption that Osiander-like skeptism was held by Copernicus did have an effect on those who were not well acquainted with the *Revolutions* which of course was most people: it was an advanced book in astronomy, not light reading. For example, Cardinal Robert Bellarmine, who was involved in Galileo's controversy with the Catholic Church, wrote in 1615 "it seems to me that your Reverence and Signor Galileo act prudently when you content yourselves with speaking hypothetically and not absolutely, as I have always understood that Copernicus spoke" (Ko454). Bellarmine means that he thought that Copernicus was only introducing calculational hypotheses, not physical ones. It needs to be said that Bellarmine was not an ignorant man; he was one of the leading theologians of his time (Fa120). We can conclude that Osiander's preface may have had some of the intended effect. Whether that effect was desirable is hard to say.

6.2.2 De Revolutionibus Orbium Coelestium

We shall only briefly discuss the *Revolutions* itself. Its fundamental hypothesis is heliocentrism. What drove Copernicus to this radical, although not entirely new, idea? His own words:

> I was impelled to consider a different system of deducing the motions of the universe's spheres for no other reason than the realization that astronomers do not agree among themselves in their investigations. ... For although those who put their faith in homocentrics (geocentric spheres) showed that some nonuniform motions could be compounded in this way, nevertheless by this means they were unable to obtain any incontrovertible result in absolute agreement with the phenomena. On the other hand, those who devised the eccentrics seem thereby in large measure to have solved the problem of the apparent motions with appropriate calculations. But meanwhile they introduced a good many ideas which apparently contradict the first principles of uniform motion. Nor could they elicit or deduce from the eccentrics the principal consideration, that is, the structure of the universe and the true symmetry of its parts. On the contrary, their experience was just like some one taking from various places hands, feet, a head, and other pieces; since these fragments would not belong to one another at all, a monster rather than a man would be put together from them (quoted from Ro4).

Two parenthetical

It is an advanced book, aimed at astronomers, not at the general public.

6.3 Tycho Brahe

6.4 Johannes Kepler

Johannes Kepler (1571–1630) was born in Weil der Stadt (then simply Weil), a small Swabian free imperial city in south-west Germany surrounded by the duchy of Württemberg (Ca29, 32, 358). His family was of the urban craftsmen class though they claimed remote noble ancestry (Ca30). Kepler's grandfather Sebald had served as mayor of of Weil der Stadt (Ca29), but his father Heinrich led a wandering and unprofitable life and abandoned his family finally in 1588 (Ca35–36). Kepler was brought up in the Lutheran denomination and adhered to that denomination all of his life. Given the religious persecutions and wars of his times, it is not surprising Kepler's religion had a significant effect on his life's course.

In 1589, Kepler entered the University of Tübingen in Würtemberg financially supported by scholarships and began training for a Lutheran clerical career (Ca41–43). He took a master of arts degree in 1591 and entered the theology faculty to obtain an advanced degree (Ca43; Ko240). Shortly thereafter the senate of the university supported the renewal of his scholarship with prescient words:

> Because the above-mentioned Kepler has such a superior and magnificent mind that something special may be expected of him, we wish, on our part, to continue to that Kepler his stipend, as he requests, also because of his special learning and ability (quoted from Ca44).

The senate and Kepler probably both expected his achievements to be in theology.

Kepler attended the lectures of Michael Mästlin, the professor of mathematics and astronomy. Publicly Mästlin adhered to the Ptolemaic system because of Copernican theory was proscibed by his colleagues as contrary to scripture. Privately he taught Copernicanism to a select few among them Kepler (Ca46). Kepler was enthused and not constrained:

> Already in Tübingen when I followed attentively the instruction of the famous Magister Michael Mästlin, I perceived how clumsy in many respects is the hitherto customary notion

of the structure of the universe. Hence I was so very delighted by Copernicus, whom my teacher very often mentioned in his lectures, that I not only repeatedly advocated his views in the disputations of the candidates, but also made a careful disputation about the thesis that the first motion (the revolution of the heaven of the fixed stars) results from the rotation of the earth. I already set to work also to ascribe to the earth on physical, or, if one prefers, metaphysical, grounds the motion of the sun, as Copernicus does on mathematical grounds. For this purpose I have by degrees—partly out of Maestlin's lecture, partly out of myself—collected all the mathematical advantages which Copernicus has over Ptolemy (quoted from Ca46–47).

6.4.??? Rosenkrantz and Kepler

As an aside (and rightly so), Kepler in 1600 during his first period with Tycho met with Frederick Rosenkrantz, Tycho's third cousin (Th427–429; Ca107–108). Since Kepler travelled briefly with Rosenkrantz, they must have communed at least a little. Like Tycho, Rosenkrantz was of the high Danish nobility and was learned. Tycho reports that he knew Latin, Italian, French, Spanish, and German. One supposes he also knew Danish. Ironically, English was not mentioned.

Rosenkrantz's good fortune had been checked, however. He had gotten a Danish lady-in-waiting, Rigborg Brockenhuus, pregnant. This seems a small sin measured against the centuries. But as Hamlet remarks "conception is a blessing; but not as your daughter may conceive:—friend, look to't" (Shakespeare 1979, vol 3, p. 351, act II, scene II). The wrath of Reformation Lutheran Denmark fell upon the lovers. Rigborg was sentenced to house arrest for life. Rosenkrantz was banished and enjoined to go on a campaign against the Turks. In a medieval manner, he had been effectively ordered to take the cross and go on a crusade as a penance. It was on the "road to Jerusalem" that he paid a visit to Tycho in April, 1600. Rosenkrantz died (1600????) soon after his visit in attempting to stop two dueling comrades.

In 1592, in an earlier adventure (misadventure?) Rosenkrantz and another remote cousin Knud Henriksen Gyldenstierne were part of a Danish legation to England. Somehow they caught the attention of William Shakespeare—in some mighty pageant of equestrian nobility? or quaffing ale in the Mermaid Tavern? In any case, Shakespeare cast them in bit parts in *Hamlet*. In the 20th century, Tom Stoppard gave them starring roles in *Rosenkrantz and Guildernstern are Dead.*

Given the tragedy of his life, it is fitting that Rosenkrantz became an immortal. And it is wonderful to discover the missing link between those geniuses of their age: Shakespeare and Kepler.

6.4.??? The End of Kepler

Amidst the collapse of towns, provinces, and countries of old and new generations, in the fear of barbaric raids, of the violent destruction of hearth and home, I see myself obliged, a disciple of Mars though not a youthful one, to hire printers without betraying my fear. With the help of God I shall indeed bring this work to an end, in soldierly fashion, giving my orders with bold defiance and leaving the worry about my funeral to the morrow (quoted from Ko420; see also Ca348).

Kepler died about noon November 17 1631 in Regensburg and was buried in the Protestant cemetery of St. Peter on November 17 (Ca358). The cemetery was demolished within a few years because of the necessities and events of the the Thirty Years' War; Kepler's grave site became lost (Ca360–361). Some words from one Kepler's last letters seem a better epitaph than the one he chose himself (Ca359):

Sagan in Seliseia, in my own printing press, 6 November 1629.

When the storm rages and the state is threatened by shipwreck, we can do nothing more noble than to lower the anchor of our peaceful studies into the ground of eternity (quoted from Ko427; see also Ca349).

6.5 Galileo

6.6 Isaac Newton

Isaac Newton was born in Woolsthorpe, Lincolnshire on Christmas day in 1642 according to the Julian calendar still being used in England (No366). The date was 1643 January 4 according to Gregorian calendar then at use throughout Catholic Europe and now universal (Ab49–50) Galileo had died only 361 days before in his 78th year??? (Ab49–50). He was born posthumous and was raised in early years by his grandmother; his mother lived with his stepfather to whom Newton was not close. Newton was schooled in Grantham and matriculated at Trinity College, Cambridge in 1661. We know that he studied the results of Kepler, Galileo, Descartes, and the mathematician John Wallis (1616–1703: Bo416), and was aware of Kepler's 3rd law and studied methods for finding planetary positions (Ha15, 55; No367).

Newton made many of his most brilliant discoveries early on, though many of them were not published until much later. In Newton's words on his early work:

> All this was in the two plague years 1665 & 1666 for in those days I was in the prime of my age for invention and minded Mathematics and Philosophy more than at any time since (quoted from Ha20).

For much of the plague period, the university was closed and Newton had retreated to Woolsthorpe (Ha31).

In 1669, he succeeded to Isaac Barrow as Lucasian professor of mathematics at Cambridge and was elected in 1672??? (Ha118) to the Royal Society, whose meetings in London he often attended. Newton left Cambridge for London permanently in 1696 when he was made Warden of the Mint. He was elected president of the Royal Society in 1703 (Ha306) and died in 1727.

Newton's scientific reputation was unparalleled in his own day and he remains one of the most esteemed scientists of history. In 1665–1666, he invented elements of calculus, but only revealed a little of his methods in the *Principia* (see § 6.1) and fuller accounts were not published until after 1700 (Bo432-433; Ha250, 264–265). Gottfried Wilhelm Leibniz (1646–1716) indepedently invented calculus and published his discoveries in 1684 (Bo438; Ha251, 258). There was a bitter priority dispute, but both persons are justly credited. (The basic elements of most of our modern calculus nomenclature and notation descend from Leibniz [Bo441].) In 1704, Newton published his *Opticks* in which he set forth his theory of light and the results to be derived from it (Ha279). This book was highly regarded by his contemporaries and their immediate followers for its very experimental character (Ha291–293). Newton's greatest work was in the physics of motion and gravity. We will now turn to these topics.

6.6.1 Newton's Laws of Motion

Newton's greatest book, first published in 1687, was his *Principia Mathematica Philosophiae Naturalis* (*Mathematical Principles of Natural Philosophy*) (No366), usually just called the *Principia* (with the 'c' pronounced 'ch'). Certainly, the *Principia* is one of the last great books written in Latin, the medieval and Renaissance language of international scholarship. In this book, Newton sets forth his three laws of motion, his law of universal gravitation, and applies these laws to many terrestrial and celestial phenomena. The *Principia* is a difficult, advanced mathematical book. The elegant, short proofs devised in later times were not available to Newton. The proofs he gave were based on a geometrical calculus of his own invention (Ha213). He already had discovered algebraic calculus (what we just call calculus) before 1671, but preferred not use it. Newton late in life, however, claimed to have originally proven his results using (algebraic) calculus and later for publication devised the geometrical proofs. This claim seems to have been made to buttress his position in the calculus priority dispute with Leibniz (Ha212–213).

The three laws of motion are:

1. The Inertial Law: A body at rest or in uniform, straight-line motion will stay in its state

of rest or motion unless acted on by a net outside force.

2. $\vec{F} = m\vec{a}$: The net force $\vec{F}$ acting on a body of mass m causes an acceleration $\vec{a}$.

3. The Reaction Law: For every force, there is an equal and opposite force.

In antique terms, the inertial law defines natural motion. However, there are questions as to how to define a straight line, time (so that you can tell that motion is uniform), and force. The straight line question is solvable in Euclidean geometry which Newton assumed to hold in the real world. The time question, Newton acknowledged to be difficult: was there any truly equable motion that could be used to define time. He assumed that mechanical clock time and celestial time (which equalled each other within error) sufficed. The nature of force we discuss now under the heading of the 2nd law.

The 2nd law, usually referred to simply as $F = ma$, tells what effect a net force will have. It is a law relating two vector quantities (i.e., quantities that have size and direction), force and acceleration, and a scalar quantity (i.e., a quantity having only size), mass. The general discussion of vectors and scalars is beyond our scope, but I assume they are understood—in a general sense. Straight-line acceleration is easy to understand if one allows that time and distance can be measured. However, a change in velocity (which is also formally a vector) occurs with changes in direction as well as in changes in magnitude; changes in direction are also accelerations in Newtonian physics.

Force requires an independent definition: $F = ma$ does not define force, it tells what a net force can do. If you measure an acceleration, you know the size and direction of the net force, but not its cause or how it acts generally. Newton came up with one outstanding example of a force defintion: is universal gravitation law to be discussed in the § 6.6.2. In modern physics only four fundamental forces are known: the gravitational, electromagnetic, strong nuclear, and weak nuclear forces. The two nuclear forces are intrinsically quantum mechanical and cannot be used reasonably in $F = ma$. The macroscopic electromagnetic force can, however. So can those nonfundamental forces derived from the electromagnetic force: e.g., frictional, fluid resistive, spring, pressure, elastic, and animal body forces. Suitable for prescriptions for most simple nonfundamental forces are known such as Hooke's law for the spring force. These prescriptions together with the third law allow one to calculate forces without observing acceleration. Thus one can calculate the balanced forces acting on an object at rest or the component forces that make up a net force that causes an acceleration.

Mass, or more exactly inertial mass, is a measure of resistance to acceleration (i.e., a measure of inertia). If one has a known force, one can apply that force to various bodies and determine their mass by observing their acceleration. One can also measure (inertial) mass by a measurement of gravitational mass as we will discuss in § 6.6.2.

One should note that strictly speaking the 1st law is redundant. It can be derived from the 2nd law; it is the special case of the 2nd law where $\vec{F} = \vec{0}$.

Another special case of the 2nd law is that of uniform circular motion. By our vector definition of velocity, such motion is accelerated motion and thus requires a net force. Any one who has turned a corner in a car knows that a force is necessary to cause a departure from straight line motion. Something which is not obvious, but follows from a vector analysis, is that the the acceleration and therefore the net force required for circular motion must be pointed toward the center of circular path. Such a center-pointing force is called a centripetal force. The $F = ma$ law in the case of uniform circular motion can be rewritten as

$$F = m\frac{v^2}{r} , \tag{6.1}$$

where we have dropped the vector notation since the direction of the force is uniquely specified, r is the radius of the circle, and v is the magnitude of the velocity of the circling mass (i.e., its speed). We note that the force and velocity are perpendicular to each other. The mass keeps trying to go a straight line and is forced into a curved path by the the centripetal force. Note also that equation (6.1) is not a prescription for a force, but tells what the centripetal force must be to cause uniform circular motion.

The 3rd law is clear enough, except to remark that the equal and opposite forces are acting on different bodies; if this were not so all forces would cancel and no acceleration would occur. The 3rd law implies conservation of momentum: a topic we will not deal with further here.

6.6.2 Universal Gravitation

Gravity in the sense that bodies fall to the Earth has been known to humankind as long as we have have been. Newton did not discover that kind of gravity. What he discovered was quantitative mathematical law of universal gravitation; this law postulated gravitation as a force generated by all matter and that gravitation was active throughout the universe. Newton himself alleged that his interest in gravity began with a fall of an apple while Lincolnshire during the in the plague years:

> It was occasion'd by the fall of an apple, as he sat in a contemplative mood. Why should that apple always descent perpendicularly to the gound, thought he to him self. Why should it not go sideways or upwards, but contantly to the earths centre? Assuredly, the reason is, that the earth draws it. There must be a drawing power in matter: ... If matter thus draws matter, it mustbe in proportion to its quantity. Therefore the apple draws the earth, as well as the earth draws the apple. [And thus] there is a power, like that we here call gravity, which extends its self thro' the universe (quoted from Ha54).

Actually, it is now believed Newton was antedating his clear conception of universal gravitation in the above passage. The clear conception was probably not present until the the mid-1680s after the Newton's thinking had been stimulated by his scientific rival Robert Hooke (Ha62, 64). Newton's own dating of the discovery of the law of universal gravitation to 1666 was probably an attempt put his priority vis-à-vis Hooke beyond question or dilution (Ha64, ????).

The modern mathematical form of the gravitation law is a vector law giving the force between two point bodies:

$$\vec{F}_{1,2} = -GM_1M_2\ r_{1,2}^2\hat{r}_{1,2}\ , \tag{6.2}$$

where $\vec{F}_{1,2}$ is the force on body 2 due to body 1, G is the universal gravitational constant, M_1 and M_2 are the gravitational masses of the bodies, $r_{1,2}$ is the distance between the masses, and $\hat{r}_{1,2}$ is a unit vector pointing from body 1 to body 2. The negative sign in the gravitation law shows that the force is attractive. By symmetry, one can see that the force on body 1 due to body 2 is equal and opposite to the force on body 2 due to body 1; this is consistent with the third law of motion. We note that the gravitational force is proportional to the two masses and is an inverse-square law in the distance separating the masses.

To find the forces between two finite bodies, one does a vector summation over the forces between all their microscopic parts: in the terminology of calculus, between all differential pieces of the two bodies. For spherically symmetric bodies that do not overlap, equation (6.1) is, in fact, the result of such a summation. If this simple result for finite bodies had not existed, then the gravitation law would have been much more difficult to deduce. Newton in fact worked backward from the spherical bodies case to the microscopic gravitation law????? (Ha?????).

The gravitational masses that appear in gravitational law are analogous to the electrical charges that appear in the Coulomb law of electrostatics: they are gravitational charges. There is no reason in Newtonian physics why gravitational mass and inertial mass (which appears in $F = ma$) should be the same, but empirically they are. Newton??? and later to experimenters to the present day have attempted to find a distinction and have failed to do so (??; Ad5). Einstein, not believing in a coincidental sameness of the two kinds of mass, asserted that the sameness implied a fundamental equivalence between gravitation and inertia (Ad4–5). He called this fundamental equivalence the equivalence principle and made it a primary axiom from which he derived general relativity (see § 8.3).

As an illustration of the sameness of gravitational and inertial mass let us consider the gravitational acceleration near the surface of the Earth. Let the gravitational mass of the Earth be M, and the gravitational and inertial masses of a small body be $m_{\rm e}$ and $m_{\rm g}$, respectively. Dropping vector notation since we know that the gravitational force and acceleration are both directed to very nearly to the center of the Earth, we find from $F = ma$ and the gravitation law that

$$a = \frac{GM}{R_\oplus^2}\frac{m_e}{m_g} = g\frac{m_e}{m_g} \; , \tag{6.3}$$

where $R_\oplus$ is the Earth's radius and g is the traditional simple for the combination of factors $GM/R_\oplus^2$. We see that if m_e and m_g are not always in a constant proportion to each other, then not all bodies will accelerate at the same rate under the force of gravity. Ever since Galileo, it had been found that all bodies do accelerate at the same rate under the force of gravity (neglecting air resistance), and so gravitational and inertial mass must be proportional. The units of the masses are chosen so that the proportionality constant is 1. The g quantity is thus the gravitational acceleration at the Earth's surface. The mean value of g at sea level is $980.6\,\mathrm{cm\,s^{-1}}$. This quantity varies from about $978\,\mathrm{cm\,s^{-1}}$ at the equator $983\,\mathrm{cm\,s^{-1}}$ at the poles due to the asphericity of th Earth and the centrifugal force (which is the subject we will not treat).

Newton could only determine the gravitational constant G very crudely given the uncertainties in his day in measuring g, the mass of the Earth, and the radius of Earth. The modern G value is $6.67259 \pm 0.00085 \times 10^{-8}$ in CGS (gram-centimeter-second) units. Note that G is still known accurately to only 4 digits; this makes it one of the most poorly determined of fundamental constants. The reason for this is that the gravitational force between laboratory size objects is so minute that measurement of this force is extremely difficult.

Gravitation is in fact the weakest of the four fundamental forces over short ranges. Nuclear forces dominate the structure of nuclei, but are very short range and have no effect a few fermis (a few times $10^{-13}\,\mathrm{cm}$) from their sources, protons and neutrons. The electromagnetic force is longer range and determines the structure of solid objects from atoms to, at least in part, asteroids. The electrostatic force component of the electromagnetic force is the main one. It is an inverse-square force like gravity, and the electrostatic force between charged fundamental particles or even charged atoms is much stronger than the gravitational force between those particles. However, there are positive and negative charges that have strong tendency neutralize each other over macroscopic distance scales. Large charge polarizations can build up of course as in thunder clouds: the sudden discharge is lightening of course. On atomic scales, quantum mechanical effects prevent total neutralization; thus a proton and an electron ordinarily do not collapse into each other to form a neutral particle. The atomic range electromagnetic forces left over after near neutralization hold solid objects together in a linked fashion. One ion, atom, or molecule holds its neighbors which hold their neighbors and so on.

Gravity, unlike electromagnetism, has only one kind of charge, mass, and mass is always attracts other mass. Thus if mass accumulates in large lumps, gravity becomes very strong and only strong electromagnetic or quantum mechanical pressure can resist gravity on the large scale. Such pressure forces have no resistance to shear forces (which are those forces that do not try to compress the material). Thus planets and stars are forced into nearly spherical shapes: gravitation acts as shear force on any nonspherical perturbation. Liquids and gases in general have no resistance to shear forces are given a macroscopic structure by gravity and pressure forces in the absence of kinetic energy and macroscopic electromagnetic forces. (This is not quite true for some cohesional forces can effect shape. Thus small water on small scales tends to clump into drops.) In nonstatic, nonelectromagnetic situations, the shapes of liquid and gases are determined by the interaction of gravity, pressure, and kinetic energy effects.

It happens that stars considered as point bodies behave somewhat like a very dilute gas. Thus, star systems (including gas and dust components), it is believed, are given their structure by gravity and kinetic energy. The uncertainty in last statement is because it is possible that some other large scale force is acting that we have not identified. Assuming that there is no such mystery force, star clusters, galaxies, and galaxy clusters are all given their structure by gravity and kinetic energy. Modern cosmology likewise holds that the structure of the universe is determined by gravity and kinetic energy (see Chapt. 8).

6.6.3 Gravity and Kepler's Laws

One of the great triumphs of Newtonian physics was that Kepler's three laws could be derived. Here we will only focus on the third law and only only for the special case of uniform circular motion.

Recall that the uniform circular motion form of $F = ma$ is given by

$$F = m\frac{v^2}{r} \ , \tag{6.4}$$

where F is the centripetal force. If we now put a large body of mass M at the center of the circle and demand that gravity supply the centripetal force, then we have

$$\frac{GMm}{r^2} = m\frac{v^2}{r} \ . \tag{6.5}$$

Recalling that the speed of uniform circular motion is given by

$$v = \frac{2\pi r}{P} \ , \tag{6.6}$$

where P is the period, one immediately finds that

$$p^2 = \frac{4\pi^2}{GM}r^3 \ . \tag{6.7}$$

Equation (6.7) is just Kepler's 3rd law for the case of uniform circular motion.

To return to history. Newton had obtained the centripetal force law as early as 1664–1665 (Ha57–59). Now Newton had known of Kepler's 3rd law since 1661 or 1662 from Thomas Streete's *Astronomia Carolina* (1661): his grasp of its importance seems to have been unusual at that time (Ha62). In a document of perhaps 1669–1670 (Ha62–63), he had put the centripetal force law and Kepler's 3rd law together, reversing the order of the derivation above, to come to an understanding that the acceleration of the planets to the Sun (approximating the orbits as circular) was proportional to the inverse-square of their distance from the Sun (Ha62, 406). But Newton does not make gravitation by the Sun the cause of this acceleration. His physical theory is still that of Descartes: vortices in the ether are the cause of the circular motion (Ha63).

In the same document document Newton shows that the ratio of the centripetal acceleration of the Moon to the gravitational acceleration on the Earth is roughly equal to $(R_\oplus/d_{\rm Mo})^{ApJstyle2}$, where $R_\oplus$ is the Earth's radius and $d_{\rm Mo}$ is the distance to the Moon. Newton would later claim that he had "pretty nearly" confirmed an inverse-square law for gravitation by this result and that the result dated back to 1666 (Ha64). This is does not seem to be the case, the document is probably from later (probably 1669–1670 as mentioned above) and Newton was did not yet have the idea of universal gravitation (Ha64).

In 1679–1680, Robert Hooke, the Secretary of the Royal Society (Ha157) and a man whom earlier and later would be on bad terms with Newton, corresponded with Newton concerning motion and forces (Ha159, 202ff, 207). Hooke asserted his belief that the motion of a body under an inverse-square law force would be an ellipse and that the Sun's force exerted on the planets was an inverse-square law force (Ha204–205). Hooke was no mathematician and could not proof this result; how he arrived at it is not clear, but probably from a qualitative argument (Ha204). Newton did not answer Hooke's crucial letter (Ha204), but he was stimulated to prove mathematically that the assertion was true immediately after the letter (January 1680) (Ha206). Characteristically, Newton put the proof aside.

Newton never published any acknowledgement of Hooke's insight (Ha205). Privately, he admitted that Hooke had stimulated his proof (Ha206), but he also contended privately that both he (before 1673) and Christopher Wren (1677) had already perceived that gravity had inverse-square law (Ha205, 425). Without clear documentation prior to Hooke's assertion of an inverse-square law for gravity (at least of the Sun's gravity), there will always be questions about priority. Given Newton's possessiveness of his discoveries, it is possible that he later over-interpreted his own earlier understanding. Hooke was likewise possessive of his discoveries and speculations, and would claim

Newton had stolen the idea of the inverse-square law of gravity. Certainly, Newton should have acknowledged Hooke in print. However, a key pointed emphasized by Newton and admitted by Hooke was that Hooke, who was no mathematician, was incapable of deriving the consequences of the inverse-square and Newton, the supreme mathematical physicist of his day, was (Ha158–159, 205, 206).

But why did the capable Newton put aside his first proof of elliptical orbits following from the inverse-square law. Hall argues that Newton simply rejected the inverse-square law as unphysical; that Newton still believed that the in Cartesian vortices where bombarding physical particles drove the planets and that this complex fluid motion could not be reduced to a simple bodiless force (Ha207). This may well be the true reason for Newton putting no great importance on his proof. But why did he not publish it just as a mathematical demonstration? Likewise why did he delay for decades in publishing his version of calculus? I would suggest that Newton's possessiveness of his discoveries is one key. As long as a discovery was known to Newton alone, then it was entirely his own. Newton was slow in realizing that his discoveries could be rediscovered. Thus, Leibniz independently invented calculus and published his version in 1684 and 1686 while Newton's calculus, that had begun to form in 1665 (Ha35) and been written up in a manuscript 1669 (Ha82) remained almost unknown. It would take a bitter and unfair dispute to establish Newton's priority in the years after 1699 (Ha257ff, esp. 262, 268).

A second possible reason for Newton's slowness to publish was simply his absorption in multiple fields of research. Solutions adequate for Newton's own understanding could be left languishing while Newton restlessly took up another problem. Besides physics and mathematics, Newton was engaged in experimental chemical and alchemical studies at least through 1669–1695 (Ha179ff), spent many years studying optics (Ha41ff, 68ff, 97ff), probably pursued world and Biblical chronological problems from his undergraduate days (Ha340), and in Biblical prophecies from 1675 at least (Ha372ff). In the later two fields Newton's achievements are of rather small significance to moderns.

The next issue is what then caused Newton to put aside his reluctance to publish (from whatever its cause) and multiple projects to work out and publish the *Principia*. Certainly, in this case encouragement by friends and a surge of inspiration impelled Newton. Edmond Halley visited Newton in August 1684 and asked if him what would be the shape of an orbit be under the influence of an inverse-square force (Ha207). Halley, Wren, and Hooke had been unable to solve this problem. Newton answered that it was an ellipse as he had known since 1680, but he found that he had lost the proof. After Halley left Newton went into and storm of creation, encouraged by Halley and the Royal Society (Ha209), that resulted in the *Principia*: a complete manuscript in the spring of 1686 (Ha215) and publication in July 1687.

In creating the *Principia* Newton abandoned Cartesianism; it is not known how or when (Ha207). He no longer required all motion to be caused by the action of body on body. Action at a distance, the force of attraction of gravity was allowed as effectively a fundamental physical fact. In a famous remark in the General Scholium of the *Principia*, Newton declares that he "feigns no hypotheses???" (????). In the modern sense, this is not true: the laws of the *Principia* are certainly all hypotheses or theories or axioms. But what Newton meant was that he would not satisfy Cartesian philosophy by trying to explain gravity in terms of the action of particles on particles if the explanation added nothing to the predictive power of his necessary axioms.

The quantitative agreement of his results within experimental error with observations was so good that Newton's laws swept all before them. He had established a new

8. The Universe of the 20th Century

8.1 The Discovery of the Extragalactic Universe

8.2 The Expanding Universe

8.3 Einstein, General Relativity, and the Einstein Universe

8.3.1 Einstein and Relativity

Albert Einstein (1879–1955) was born in Ulm???? Germany. For a time he and his family lived in Munich where there is an Einsteinstraße on which street they presumably lived. When I lived in Munich (1988–1989) I never went looking for the Einstein mansion, assuming one day that I would just find myself there; but that day never came. Einstein disliked the regimented, nationalistic ethos of pre-World War I Germany, its antisemitism, and all authoritarianism. Everyone as a favorite Einstein quote. Mine, quoting from memory, is approximately:

> "When I was young, I despised all authority and I have been punished for it. Society has made me into an authority myself.

When a teenager???, he left home to live in Switzerland. He took his university degree there and became a Swiss citizen: he was always very proud to be Swiss????. The Swiss are, of course, a superior form of life.

Following his graduation he settled in Bern??? and obtained a position as a patent clerk in the Swiss patent office. Besides settling down to being a family man with wife and children????, he also carried on his interest in his free time in physics. Originally, he had been most interested in experimental physics???, but without opportunities for laboratory work, he perforce became a theoretician. In 1905, his Anno Mirabilis, he published his famous papers on Browning motion and the photoelectric, and his first paper on special relativity. He also obtained a his Ph.D. that year with an unmemorable dissertation????.

Special relativity was motivated by a need to perfect electromagnetism of James Clerk Maxwell (1831–1879), summarized in Maxwell's equations. These equations admit solutions which are free electromagnetic wave solutions with a phase velocity (derived entirely from electric and magnetic effects) that within uncertainty is consistent with the speed of light. With this result the idea quickly arose that light was a form of electromagnetic waves or electromagnetic radiation. Heinrich Rudolf Hertz (1857–1894) confirmed this idea by producing and detecting long-wave electromagnetic radiation in the laboratory in 1888. This long-wave electromagnetic radiation is, of course, radio. But if light and other electromagnetic radiation are wave phenomena, what is the medium in which the waves oscillate? This otherwise unknown medium was given the name the ether (a revival of the Greek term for different purpose). The ether was supposed to pervade space and be at rest in the reference frame of absolute space. Maxwell's equations are in fact not invariant under Galilean transformations and it was thought that they could only apply in one special reference frame, the frame of the ether. This is unlike Newtonian dynamics which apply in all inertial frames. Maxwell himself thought his equations would need to be modified to account for motion, in particular the motion of the Earth. The fact that his equations seemed to work to a high degree of accuracy in the moving frame of the Earth suggested the correction would be small.

In 1879, the year of his death, Maxwell proposed an experiment using interference to measure varying speed of light due to the Earth's motion through the ether (Hec7). Albert A. Michelson (1852–1931) and Edward W. Morley (1838-1923) carried out this experiment (reported in 1887) and obtained one of the momentous null results in history: the speed of light did not within their uncertainty depend on the motion of the Earth. This null result caused a flurry of complicated theorizing about electrodynamic effects by physicists among whom the most prominent were Hendrik Antoon Lorentz (1853–1928), ???? Fitzgerald (????), and Jules Henri Poincaré (1854–1912). It was shown first by Voight???? then by Lorentz that the Maxwell equations were invariant under a special set of transformations now called the Lorentz transformations. But the physical significance of the Lorentz transformations was not satisfactorially understood.

Einstein's papers on special relativity beginning in 1905 cleared the difficulties away. His two postulates were: (1) the speed of light (in vacuum) is the same for all observers no matter how they are moving, and (2) the laws of physics are the same in all inertial reference frames. From these two postulates one obtains the Lorentz transformations under which Maxwell's equations are invariant. However, Lorentz transformations must also apply to material kinematics. Thus, the Galilean transformations and Newton's law of motion are not valid except in the low velocity limit (i.e., for velocities much less than the speed of light). From the Lorentz transformations, strange effects like the twins paradox and the Fitzgerald contraction are obtained. These effects are only noticeable at high velocities, but they have been well verified. The effects testify to the interdependence of the three spatial dimensions and the time dimensions: the unified term for space and time is space-time. But one should note there is an asymmetry: the spatial dimensions and time are not on a equal footing. At least in one's own frame, one can move in time alone (but only one way), but, unless one is a light beam, one cannot move orthogonally to time. Time is the strange dimension which, at least in the macroscopic world, does not tarry or return.

From special relativity also follows the equation

$$E = mc^2 \; , \tag{8.1}$$

where E is energy, m is mass, and c is the velocity of light. What this means is that energy and mass are really the same thing: mass-energy. However, there are different forms mass-energy: electromagnetic radiation, electromagnetic potential energy, gravitational potential energy, kinetic energy, and rest mass. There is not always a convenient way of transforming one form of mass-energy into another. Thus, any old gram of matter cannot simply be converted into enough electric energy to power a 100-watt light bulb for 3×10^5 years, alas. Nature is too stable to allow this: this stability is understood in terms quantum mechanical conservation laws. Perhaps, it is a good thing that all the protons of our bodies cannot "evaporate" of a sudden in a smallish bang.

Although, more correct than Newtonian physics, special relativity is less complete. It does not deal with accelerations???? or forces. Thus, Einstein undertook the task of generalizing his theory: first for gravitation and accelerations (in which he succeeded with general relativity) and later for electromagnetic forces (in which he essentially failed with his unified field theories).

8.3.2 General Relativity

Einstein's theory of gravity and motion under the influence of gravity is general relativity. This theory was given its classical form in 1915. By that date Einstein had given up his promising career in the Swiss civil service, had held a professorship in Prague, and was established at the Prussian Academy of Sciences in Berlin. general relativity states that mass-energy modifies the geometry of space-time (which in the absence of mass-energy is Euclidean in the three spatial dimensions). Movement in this space is along geodesics (the generalization of straight lines to general Riemannian geometries). Thus, the force of gravity explicitly disappears from Einstein's formulation.

But if mass-energy determines space-time geometry which then determines how mass-energy moves, then in general there is a "circular" situation which must be self-consistent. Nature has no trouble: it is always consistent with itself and solves all problems by complete analog computation. Most physical laws, including general gelativity, are expressed (by us) as differential equations. What happens in any given situation must be determined (by us) by solving the differential equations plus boundary conditions. Newton's $F = ma$ is a differential equation; Maxwell's equations are differential equations. The differential equations of general relativity are Einstein's field equations:

$$R_{ij} - \frac{1}{2} g_{ij} R - \Lambda g_{ij} = \frac{8\pi G}{c^4} T_{ij} \tag{8.2}$$

(e.g., Col26), where R_{ij} is the tensor, R is the Ricci scalar, g_{ij} is the metric tensor, Λ is the cosmological constant, G is gravitational constant (just as in Newton's gravitational laws), c is the speed of light, and T_{ij} is the energy-momentum tensor. The tensors in this case can be represented by 4×4 symmetric matrices. Because of the symmetry there are only 10 independent equations to solve in general—it's enough. For present purposes it suffices to state that the left-hand side of equation (8.2) represents space-time geometry and the right-hand side represents mass-energy. Thus, the two interacting entities are coupled by the equations.

The cosmological constant Λ does not appear in standard general relativity (i.e., Λ is set to zero). It is a free parameter introduced by Einstein for cosmological reasons (see § 8.3.3) and given a value by fitting. Free parameters in theories are considered undesirable. The standard formulation of general relativity is considered particularly elegant because the constants (i.e., G and c) that appear in it were determined prior to the advent of general relativity. An absolute distinction between Λ and the other constants, however, seems a bit artificial to me. All fundamental constants are after all parameters of nature which we cannot (yet) derive. The effect of the cosmological constant, at least in cosmological modelling, is something like a uniform density throughout the universe (Col18). The sign of this density-like effect is the opposite of the sign of Λ. For $\Lambda > 0$, there is a uniform repulsion throughout space: a uniform antigravity-like force. For $\Lambda < 0$, there is a uniform attraction throughout space: a uniform gravity-like force. The density equivalent of the cosmological constant is given by

$$\rho_\Lambda = -\frac{\Lambda c^2}{4\pi G} \ . \tag{8.3}$$

It has to be emphasized that the cosmological constant does not give rise to a force just like the gravitational (or antigravitational) force of a uniform density ρ_Λ, but only to a force somewhat like a ρ_Λ-density force.

In Table 7.1, we show the mean densities for various astrophysical systems. Because the density in these systems varies tremedously, mean density is not a very useful quantity, except for the two universe systems where density variations occur on a distance scale too small to be important (by the assumption of the cosmological principle: see below). However, because gravity is a long range inverse-square force, the magnitude of gravitational effects in a system is crudely measured by the mean density. (This last remark can be made more definite using the Birkhoff theorem???? Col23.) If $|\rho_\Lambda|$ is much less than a system density, then the cosmological constant cannot have an important effect on that system. At present, there is no pressing reason invoke a significant $|\rho_\Lambda|$ for any astrophysical system, and thus no need for the cosmological constant. However, it is impossible to rule out a cosmological constant giving rise to $|\rho_\Lambda| \lesssim 10^{-29}\,\mathrm{g\,cm^{-3}}$. Thus there could be a cosmological constant that affects the structure of the whole universe, but not smaller structures.

Table 8.1. The radii, masses, and mean densities for various astrophysical systems

System	R	M ($M_\odot$)	ρ ($\mathrm{g\,cm^{-3}}$)
Solar system to Pluto	39.7 AU	1.001	2.27×10^{-12}
Solar system to the Oort Cloud	$\sim 10^5$ AU	1.002	$\sim 10^{-22}$
Spiral galaxies	~ 30 kpc	10^{12}	$\sim 10^{-24}$
Galaxy clusters	~ 3 Mpc	3×10^{14}	$\sim 10^{-28}$
Universe	$3000h^{-1}$ Mpc	...	$\gtrsim 0.4 \times 10^{-29}h^2$
$\Omega = 1$ Friedmann universe	$3000h^{-1}$ Mpc	...	$\gtrsim 1.879 \times 10^{-29}h^2$

Note.—As usual in physics, ρ is the symbol for density. The mass of the Solar System includes the Sun, Jupiter, Saturn, Neptune, Uranus, Earth, and Venus. The Oort Cloud radius is taken from Ze251. The data for the spiral galaxies and galaxy clusters is derived from Ze444 and Ze450, respectively. The $\Omega = 1$ Friedmann Universe is discussed in § 8.4. The radius of the universe and the $\Omega = 1$ Friedmann universe is not a real radius. It is the characteristic radius of the observable universe:

$$R_{\rm Obs} = \frac{c}{H_0} = 3000h^{-1}\,{\rm Mpc}\ . \tag{8.4}$$

If Hubble Law applied to all distances and the inverse Hubble Constant was the age of universe, then we could not see beyond $R_{\rm Obs}$ because no light from farther away could have reached us in the universe's lifetime. Since the conditions of the previous sentence are not exactly true because of general relativistic and evolutionary effects, $R_{\rm Obs}$ is only a characteristic radius for the observable universe.

The lower limit for the density of the universe, density of $\Omega = 1$ universe, and why both these values depend on the reduced Hubble Constant are discussed in § 8.4.

---

general relativity is not an extensively tested theory. The three old tests are as follows. First, in the limit of weak gravitational fields (which is almost always obtained in the solar system) general relativity should yield Newton's graviational law and his laws of motion under the effects of graviation. This general relativity does; it was, of course, constructed so by Einstein.

Second, a problem of Newtonian celestial mechanics that arose in the 19th century was that there was an inexplicable 41 arcsecond per century shift of the perihelion of Mercury's orbit on top of the herihelion shift due the perturbation of other planets (Ze136). It was hypothesized that that an undiscovered planet nearer to the Sun than Mercury was the giving rise to the extra perturbation (Ad200). This planet, prematurely named Vulcan, has never been discovered though its name has since passed into legend. In 1915, Einstein using his theory predicted a 43 arcsecond per century shift due to strong gravitational effects; this prediction agree within uncertainty with the measurement. As the nearest (discovered!) planet to the Sun, Mercury is the planet in the strongest gravitational field and thus is most affected by strong gravitational field effects.

The third old test was the bending of light beams near the Sun. Arthur S. Eddington in 1919 confirmed this prediction in 1919 (No518). Due to the difficulty of the measurements, the degree of quantitative confirmation was in dispute for many years. Currently, there is agreement to within 10 % (Ze136).

Newer tests of general relativity have appeared since 1960. In 1960, Pound and Rebka confirmed gravitational redshift, another prediction of general relativity (Ad140). Gravitational lensing of distant quasars and galaxies by foreground galaxies was discovered in the 1970s???? (Ze471). This effect, however, is only the light beam bending effect in another context.

The most dramatic new confirmation has come from the binary pulsar PSR 1913+16 discovered in 1975 by R. A. Hulse and J. H. Taylor (Shapiro & Teukolsky 1983, p. 479). The orbit of this system is decaying. The rate of decay is within small uncertainty what is predicted by general relativity. In the general relativistic explanation, the decay is due to the loss energy by gravitational radiation from the system. Gravitational radiation has never been detected directly so far, although a large project (LIGO) is afoot to try to do just that. For their discovery and continuing analysis of the binary pulsar, Hulse and Taylor were awarded the Nobel Prize in 1992/3/4????.

General relativity has dramatically passed the available tests. Furthermore rival theories of gravitation (such as the Brans-Dicke scalar-tensor theory of gravitation [Ad380]) either fail to pass or must be adjusted to give nearly the same results as general relativity. Such theories are more complicated and less elegant than general relativity, and so can be rejected by Ockham's razor. Nevertheless general relativity is substantially less well established than quantum mechanics and quantum theories of interactions (i.e., forces).

Quantum mechanics has been known since 1925/1926????, and has been the basis of all advances in molecular, atomic, nuclear, and solid state physics. The three other forces of nature (the electromagnetic, strong nuclear, and weak nuclear forces) can all be given a quantum formulation. The weak nuclear and electromagnetic force have even been successfully united in a quantum electroweak theory. All three other forces have been incorporated in the (quantum) standard model of particle physics which if naggingly incomplete is nevertheless fairly successful.

Given, the tremendous success of quantum theories, most physicists believe that gravity must have a quantum theory also. In this opinion, general relativity can be at best the macroscopic limit of a quantum theory of gravity. Just as general relativity subsumed Newtonian gravity, the quantum theory of gravity will subsume general relativity. At present, despite massive efforts, there is no solid quantum theory of gravity. The gravitational force is just too weak between microscopic particles to allow any direct experimental access to quantum gravitational effects. Of course, the general view may be wrong: perhaps, gravity is completely different from the other forces and is not at all quantum mechanical. However, history (which is all we can rely on) suggests that all physics is closely united; physicists can only (or only want to) follow that path further.

A grander dream beyond quantum gravity, is the unification of the four forces into a single theory, a unified field theory. Einstein, himself spent his last 30 years trying to construct a theory of that name, but including gravity and electromagnetism alone. Since he rejected quantum mechanics as not fundamental and ignored the two nuclear forces, he is almost universely seen has having been on the completely wrong track. However, his famous quest set the agenda for physics ever since.

The unified field theory itself, one hopes will be subsumed into TOE: the theory of everything. Such a theory would explain all the forces, all the constants, all the particles, why quantum mechanics works, and the universe as a whole: microcosm and macrocosm united just as the ancient Stoic philosophers believed it should be. Visionaries like Hawking sometimes predict that TOE is almost within sight: in 1980, he suggested TOE by 2000 (Ov359). But there is no way to know if it is nearby. Moreover, nothing guarantees reality is a monism, not a pluralism.

8.3.3 Relativistic Cosmology

Gravity, though the weakest of the four forces at short ranges (i.e., microscopic ranges) (Ze493) is the only one, effectively, that can act at long distances because of its long range inverse square law behavior. This behavior is demonstrated by Newton's gravitational law (which approximates general relativity in the weak field limit, of course) for point charges:

$$\vec{F}_{1,2} = -\frac{GM_1M_2}{r_{1,2}^2}\hat{r}_{1,2} \ , \tag{8.5}$$

where $\vec{F}_{1,2}$ is the force on particle 2 caused by particle 1, $G = 6.6732 \times 10^{-8}$ in CGS (centimeters, grams, seconds) units, M_1 and M_2 are the masses of the two particles, $r_{1,2}$ is the distance between the particles, and $\hat{r}_{1,2}$ is a unit vector pointing from particle 1 to particle 2. The electrical force is also an inverse square law and is intrinsically very strong. But the two kinds of electric charge usually neutralize each other in macroscopic behavior and virtually always neutralize each other on astronomical scales. Any deviation from neutrality immediately gives rise to a strong attractive or repulsive force which tends to restore neutrality. On atomic scale, quantum effects prevent neutralization, and so electromagnetic interactions determine material properties. Gravity, however, has only one kind of "charge", mass, and only one force direction, attraction. Thus, large gravitational forces are built up with large mass concentrations, and these forces extend over astronomical and cosmological distances.

Because of its long range reach, gravity is the force which determines astronomical and presumably cosmological order. For cosmology, Newton's theory of gravity is ambiguous. A finite array of masses spread in otherwise empty infinite space cannot exist in a stable static configuration. If one attempts to establish such an array, the masses will be drawn into a single clump held up by electromagnetic or other forces, or collapse into a singularity (if such things can exist). There are two ways of holding a finite array from collapse, and both are nonstatic. First, give the masses sufficient kinetic energy to escape their mutual gravitational attraction. The masses then will then spend eternity expanding away from each other: i.e., will escape to infinity. The second is to set them in rotation about their center of mass. In this second case, the postulate of absolute space (i.e., space endowed with some active power) must be invoked because otherwise there is no way to decide if net rotation is occurring at all.

Newton and his contemporaries do seem to have believed in infinite space???? with a stable static configuration of stars. If space is infinite and contains infinitely many stars uniformly distributed on the large scale, is a stable static configuration possible. Newton thought this might be so, but was unable to prove it (No376). However, there is another difficutly with stable static infinite distributions of stars: the darkness paradox. The first known discussion of this paradox was made by an anonymous person and was reported by Edmond Halley (1656–1743) in 1721 although not in entirely correct form (No377). The darkness paradox states that in a infinite stable static universe with an infinite uniform distribution of stars, in every direction one looks one should see a star. This means that the universe should be in thermal equilibrium: the sky should be bright as a star and the Earth should have the same temperature as a star. Since Newton was presding over the Royal Society meeting where Halley first discussed the darkness paradox, he was probably well aware of it, but perhaps too old to worry about iti: he was about 79 at the time.

In later times, it was suggested that the darkness paradox could be solved by intervening dust or gas, or by light decaying faster than $1/r^2$ when it got old (i.e., after travelling over cosmological distances). However, if energy is conserved (which is only an adequate theory after all), then in both explanations there must be some place for the lost energy to go. In the first explanation, we now know the dust would just heat up and re-radiate energy: thermal equilibrium would be established again. In the second explanation, who knows.

Because of the stability, infinity, and darkness paradox problems, cosmology was in an unsatisfactory state (in the sense that no one knew quite how to develop it????) from disposal of Aristotelean cosmology until the advent of general relativity. Because it is a theory of gravity, general relativity clearly had cosmological implications. Originally, Einstein was not particularly interested in astronomy and cosmology, but armed with general relativity he sallied forth into the heavens. The actual universe with its stars, star systems, and nebulae, then and now is to complicated to handle in detail in cosmology. Einstein and all others???? replace the actual distribution of matter with an simplified distribution of matter that interacts only through gravity. To make the problem tractable at all the what has come to be called the cosmological principle is invoked. This principle states that on large scales the universe is homogeneous and isotropic. Thus the simplified distribution of matter is usually thought of as a constant density gas filling all space.

In addition to the cosmological principle, Einstein assumed that the universe was static. He believed that a static universe was most consistent with astronomy as then understood. However, he could not find a static solution for a universe obeying the cosmological principle with standard general relativity. It was for this reason that he first invoked the cosmological constant, Λ and fine-tuned its value to obtain a static universe. This Einstein universe presented in 1917 is a finite (i.e., closed), but unbounded, universe.

The spatial geometry of this universe, determined by the matter, is 3-dimensional, but non-Euclidean. The geometry is the same a that of the surface 4-sphere: i.e., a sphere in a 4-dimensional space.

To understand this kind of space, consider the 2-dimensional geometry of the surface of an ordinary sphere. This surface is an unbounded, but finite space. Lines on the spherical surface that begin parallel will eventually meet. The shortest distance between any two points (which has the general name of geodesic in geometry) is a segment of a great circle (i.e., circle which divides the sphere into hemisphers). Over small regions the surface is approximately Euclidean. Thus minute beings on the surface might think they live on a flat plane (a 2-dimensional Euclidean space) if they do not travel far. This is, of course, the situation for early peoples on the spherical Earth.

If one takes a sphere pushes it through a plane, then two dimensional beings living on the plane who are unaware of the third dimension would see a point appear, then a growing circular disk, then a shrinking circular disk, and finally a point again that vanishes. This behavior can be described mathematically. The equation of a sphere is

$$x^2 + y^2 + z^2 = R^2 \ , \tag{8.6}$$

where R is the radius. We can re-arrange this equation to get the equation of circle with a radius R_* parameterized by z:

$$x^2 + y^2 = R_*^2 \ , \tag{8.7}$$

where

$$R_* = \sqrt{R^2 - z^2} \ . \tag{8.8}$$

Pushing the sphere through the plane is equivalent to varying z from $-R$ to R.

The 3-dimensional surface of 4-sphere surface is similar to the 2-dimensional surface of a sphere. Lines that start out parallel will eventually meet. The shortest distance between any two points is a geodesic. If one travels along a geodesic from a point A, one will eventually return to point A. If one pushes a 4-sphere through a 3-plane (i.e., a 3-dimensional Euclidean space, then the 3-dimensional beings confined to the 3-plane will see a point become a growing sphere which reaches a maximum and shrinks away to vanishing again. The continuum of spherical surfaces the 3-dimensional beings see make up the non-Euclidean 3-dimensional surface of the 4-sphere. To minute beings living on the surface of a 4-sphere, space seems to be a 3-dimensional and Euclidean.

In the Einstein universe, we are the minute beings who do not perceive the curvature of their space. One may ask about that fourth spatial dimension. Does the Einstein universe imply that it exists in a physical sense? No: the fourth dimension can be used for descriptive purposes, but there is no indication that anything exists beyond the surface of the 4-sphere. In science fiction stories, one often encounters the idea of hyperspace where faster than light travel is possible. Presumably, the authors often mean hyperspace to be off the 4-sphere surface where you can take shortcuts. It is not apparent to me that even if such extra realm existed that we could find access to it or that travelling in it would accelerate our flight plans substantially.

What about the darkness paradox in the Einstein universe? First consider the 2-dimensional surface of a sphere again. To complete the analogy to relativistic space, light on the spherical surface must travel along great circles. If one looked all around on the spherical surface, one would see all the stars twice: once when the line of sight went less than halfway around the surface and, when looking in the opposite direction, looking more than halfway around the surface. The farthest object one would see in any direction would be the back one's own head provided nothing along the way blocked the view. The number of stars one counts is finite, and one would not see a star in every direction unless one were on star. But the light beaming in all directions from any star would diverge only for first quarter of a great circle of flight, then would converge toward the star's antipode, diverge again after passing the antipode, and converge again on the star itself. Thus the incoming light on star from the star itself should equal what it was emitting. If the surface has many stars uniformly distributed, it seems to me that the sky would be star-bright on average for any observer. Thus, the darkness paradox remains in spherical surface universe. I am guessing that the same is true in the Einstein universe: one sees finite stars, but the sky should be star-bright. In a more general argument, I think a static closed universe without perfect insulating barriers must be in thermodynamic equilibrium: i.e., be all at the same temperature. Our universe is clearly not in thermodynamic equilibrium; I think the Einstein universe fails on this ground alone.

There are two other objections to the Einstein universe. First, the use of the cosmological constant. Neither, Einstein nor many others liked it (Ad431; No520). It was considered unaesthetic and ad hoc. To be more explicit, if the original standard formulation of general relativity did not account for all gravitational effects, then the real behavior of gravity might be more complex. Introducing an adjustable parameter to patch up a particular problem (even if it is a problem of the universe as whole), may be just sweeping real problems under the rug or, as Einstein found, hiding real predictions. It was not trusting his original theory and invoking the cosmological constant that lost Einstein his chance to predict the expanding universe. After the expanding universe was discovered, Einstein called the cosmological constant the biggest blunder of his life???. (Einstein probably said this many times, but he certainly said it to George Gamow, the parent of big bang cosmology [Ga44]). In defense of the cosmological constant it must be said that it is the only simple modification of general relativity available. It has never gone away, but lurks in the shadows of the minds of cosmologists. In modern inflationary cosmology (see § 8.7), a time-dependent cosmological function (which may not be not a constant in space either????) arises as a manifestation of the scalar potential(s) that drive inflation: i.e., the scalar potential(s) enter the Einstein field equations like the cosmological constant (Col134????).

The second objection is that the Einstein universe is unstable as pointed out by Lemaître in 1931 (Ad431). Any small perturbation of the Einstein universe will start it monotonically expanding or contracting. A simple example of physically unstable system is a ball placed exactly on top of a smooth hill. Placed exactly it is balanced, but the slightest perturbation causes it to accelerate away. The stable counterexample is a ball at the bottom of smooth depression. Any perturbation will set the ball oscillating about the bottom. If any dissipative forces are present, the ball will come to rest in the bottom again.

Besides not describing the actual universe, the three objections (lack of explanation for the lack of thermodynamic equilibrium, use of the ad hoc cosmological constant, and instability) make the Einstein universe a rather unsatisfactory theory. It was a useful pioneering effort, but Einstein, as he himself well knew after 1930 (No526), had really blown it.

8.4 Friedmann Cosmology

8.5 A Big Bang or Steady State Universe

In the

Table 8.2. Fundamental Constants and Planck Quantities

Quantity	Symbol	Combination of Fundamental Constants	Value
Speed of Light	c	...	$2.99792458 \times 10^{10}\,\mathrm{cm\,s^{-1}}$
Gravitational Constant	G	...	$6.6732 \times 10^{-8}\,\mathrm{cm^3\,g^{-1}\,s^{-1}}$
Planck Constant (h-bar)	$\hbar$	...	$1.05457266 \times 10^{-27}\,\mathrm{erg\,s}$
Boltzmann Constant	k_{B}	...	$1.380658 \times 10^{-16}\,\mathrm{erg\,K^{-1}}$
Planck Density	ρ_{P}	$\dfrac{c^5}{\hbar G^2}$	$5.1565 \times 10^{93}\,\mathrm{g\,cm^{-3}}$
Planck Energy	E_{P}	$m_{\mathrm{P}} c^2$	$1.9562 \times 10^{-5}\,\mathrm{ergs}$
Planck Length	ℓ_{P}	$\sqrt{\dfrac{\hbar G}{c^3}}$	$1.6161 \times 10^{-33}\,\mathrm{cm}$
Planck Mass	m_{P}	$\sqrt{\dfrac{\hbar c}{G}}$	$2.1766 \times 10^{-5}\,\mathrm{g}$
Planck Number Density	n_{P}	ℓ_{P}^{-3}	$2.3691 \times 10^{98}\,\mathrm{cm^{-3}}$
Planck Temperature	T_{P}	$\dfrac{E_{\mathrm{P}}}{k_{\mathrm{B}}}$	$1.4169 \times 10^{32}\,\mathrm{K}$
Planck Time	t_{P}	$\sqrt{\dfrac{\hbar G}{c^5}}$	$5.3908 \times 10^{-44}\,\mathrm{s}$

Note.—The Planck quantities are combinations of fundamental constants. As one goes back in time toward the singularity of the Friedmann model, the Planck Density is reached at about the time one is a Planck time from the singularity!!!!!. At the Planck Density, it is assumed that general relativity ceases to be an accurate theory and that quantum gravity must become important. Since no accepted quantum theory of gravity exists, we cannot confidently extrapolate the big bang model back before a Planck time (i.e., before $\sim 10^{-43}$).

8.6 The Large Scale Structure

8.7 Inflationary Cosmology

Despite the amazing success of big bang cosmology, there is some dissatisfaction with it. This is summarized in two questions: what happens before the Planck time (i.e., $\sim 10^{-43}$ [see § 8.5,

Table 8.2) from the singularity and what sets the initial conditions of the big bang? Since we believe a fundamental theory of gravity exists (as an acticle of faith), we should be able to extrapolate back to before the Planck time. The fundamental theory of gravity is probably a quantum theory, but being open-minded it could be something else. It is hard to see that it could be general relativity itself because quantum effects must become important at the Planck time????. Naturally even the fundamental theory of gravity may exhibit a singularity that marks the beginning of time: i.e., there is no "what before". There may also be no "what" setting the initial conditions. If so, the beginning of time and the establishment the initial conditions of the big bang may be absolute fundamental facts of nature in their own right. Creation *ex nihilo* is no less plausible than anything else at least in some philosophies. However, not being able to recognize a fundamental fact when we see one, we would like to try to regress further back in time and deeper into physics. inflationary cosmology is an attempt at such a regression.

8.7.1 The Origins

Inflationary cosmology originated in inflationary theories by A. A. Starobinsky in the former Soviet Union (Lin51) and Alan Guth in the U.S. in 1979; Guth's work was made public in 1980 (Ov245–247): Starobinsky and Guth's work was independent and they were unaware of each other. Guth coined the term inflation to describe a superfast expansion driven by a reservior of energy that could be invoked in particle physics (Ov245). In big bang cosmology, the thermonuclear big bang gives the initial impetus for the expansion which is then only decelerated by gravity for the rest of the Friedmann evolution. Guth was not originally interested in cosmology: he was a theoretical particle physicist. But particles with sufficient energies to test some of the predictions of modern particle physics are well beyond what can be obtained by any foreseeable particle accelerators. Only in the big bang were particles of such energies assumed to occur in nature. So the big bang became the experiment for particle physicists.

In his work, Guth (probably among others) found that magnetic monopoles should be as common as protons in the big bang (Ov237, Li48). A magnetic monopole is a particle with an isolated north or south magnetic pole. Magnetic monopoles are analogous to particles with positive or negative electric charge such as protons and antiprotons. However, the predicted mass of magnetic monopoles is 10^{16} times that of protons (Lin48). Magnetic monopoles have never been verifiably observed. Thus, a particle-physics cosmology must get rid of the magnetic monopoles. Guth found that by invoking an inflation epoch he could dilute the density of magnetic monopoles so that there would only be of order one in the observable universe (Ov245). (Blas Cabrera of Stanford University reported a magnetic monopole detection in 1982; another one was reported later [Wolfson & Pasachoff 1990, p. 700]. Since two is consistent with of order one, inflationary cosmology cannot yet be ruled out on the grounds of over-abundance of magnetic monopoles.) After the inflationary epoch (which is very brief: see below), the ordinary nucleosynthetic big bang occurs followed by ordinary Friedmann evolution.

8.7.2 What its Good for

Guth and others saw that the two big questions about the big bang could answered by inflation (e.g., Ov244–245, Li48-49). Inflation came before the big bang and inflation set the initial conditions of the big bang. Leaving aside the question of "before" until after, what are the problems with the initial conditions in need of explaining? We will only discuss the two most commonly mentioned problems: the flatness problem and the horizon problem.

As discussed in § 8.4, the present day value of the density parameter probably is limited by $0.2 \lesssim \Omega_0 \lesssim 2$ (Col81, 142). Recall that $\Omega_0 = 1$ gives a Euclidean (i.e., flat) universe. Recall also that $\Omega(t)$ (the time dependent density parameter), if different from 1, must always diverge from 1 in Friedmann cosmology (except after the time of maximum expansion in a closed universe) (Col43). Thus if $\Omega(t) < 1$, it decreases forever; if $\Omega(t) > 1$, it increases until the time of maximum expansion. For a Friedmann universe (without the cosmological constant) it can be shown that

$$\Omega\left(t_{\mathrm{P}}\right) \approx 1 + \left(\Omega_0 - 1\right) \times 10^{-60} , \qquad (8.9)$$

where $\Omega(t_P)$ is the density parameter Planck time (Col141). Since Ω_0 is so close to 1 now, it $\Omega(t_P)$ must have been 1 to within an additive term of $\sim 10^{-60}$. This fine-tuning of initial density parameter makes people wonder? Maybe there is some law of nature that forces $\Omega(t) = 1$ at all times. Inflation, however, offers an alternative explanation.

In inflation it turns out (in a way I do not understand) that $\Omega(t)$ of the inflationary epoch is driven to be nearly 1 to such a high accuracy that even after inflation ends and $\Omega(t)$ begins to diverge from 1 (as it must in Friedmann evolution) it remains close to 1 till the present day. The prediction of one inflationary model is that

$$|\Omega_0 - 1| \lesssim 10^{-45} \tag{8.10}$$

(Col141, 144). Equation (8.10) is, however, an average value; fluctuations on the scale of the observable universe mean that the observable universe's Ω_0 could differ from 1 by $\sim 10^{-5}$ (Col145).

The prediction that Ω_0 should be nearly is very solid. If Ω_0 differs from 1 by much more than $\sim 10^{-5}$, then inflationary cosmology would be much less attractive than it is. However, particle physicists are clever and would organize a rescue mission. They would come up with inflationary theories without having Ω_0 nearly 1. One way is, of course, to have $\Omega_{0,\text{eff}}$ nearly 1 (Col155). A cosmological constant that is the relic of the inflationary epoch would be what builds $\Omega_{0,\text{eff}}$ up to nearly 1. Nevertheless, such a rescue would be very ad hoc, and without any new attractions inflationary cosmology might fade away.

The horizon problem is essentially that the temperature of the CBR is so constant in all directions that it is hard to believe that the big bang region corresponding to the observable universe was not causally connected. (Recall from § 8.5 that the mean???? CBR temperature is 2.726 ± 0.005 K with directional varitions of $\Delta T/T \approx 10^{-5}$ [Col92-93].) By causal connection, I mean that the observable universe at some point before the recombination epoch had a characteristic size such that light could travel across it in at most the time from the beginning to the recombination epoch. Causal connection would have allowed thermodynamic equilibrium (i.e., constant temperature) to be achieved. The alternative point of view is that thermodynamic equilibrium is just a fundamental initial condition. But if one does not want to believe in such initial fine-tuning or one does not want to believe the big bang was a beginning, then causal connection looks necessary. However, in the ordinary big bang model the different parts of universe from whence the CBR now reaches us were too far apart to be causally connected. Inflation solves this simply by having the entire observable universe inflated from a domain (i.e., a patch of space) small enough to be causally connected, and thus in thermodynamic equilibrium without fine tuning. The inflation of this domain is "faster" than the speed of light. However, no real velocity is involved, rather just the creation of space????.

8.3 An Example Inflationary Model: Self-Reproducing Inflation

Inflationary models are many: e.g., old inflation, new inflation, chaotic inflation, stochastic inflation (sometimes called eternal inflation or is that infernal inflation), and extended inflation (e.g., Col151ff). Some of these models have been ruled out (e.g., old inflation [e.g., Li51]), but others are still tenable. The inflation model recently devised by Andrei Linde and colleagues (Lin) seems to be representative of the current state of the art in Inflationary cosmology: self-reproducing inflation. We will discuss this inflation theory as an example.

All inflationary theories make use of scalar fields as sources of energy which is stored in the scalar fields as a form of potential energy. Scalar fields are not always exotic. Both electricity and grativation can be described by scalar fields. The voltage of a wall outlet is a measure of the potential energy stored in the electric potential scalar field. In modern particle physics scalar fields have been used to devise theories of particle interaction that have been experimentally verified. However, particle physicists find it easy to invent other scalar fields which may exist, but are currently unamenable to direct experimental verification.

Linde postulates a small region of primordial space with a characteristic length that is of order of the Planck Length: i.e., of $\sim 10^{-33}$ cm (see § 8.5, Table 8.2). This small region is called a domain. In this domain there are one or more scalar fields of high potential energy. This high energy state is unstable due quantum fluctuations??? (intrinsically random variations in physical quantities). A quantum fluctuation dislodges the scalar field analogous to a small knock disturbing a ball at rest on top of a hill. Dislodged the scalar field can only "run downhill" and have its potential energy changed into other forms: these forms are the creation of space???? space, massive particles, and radiation. The space creation, called inflation, is much more rapid than expansion in the big bang models due to nuclear energy release. In a time of $\sim 10^{-35}$ s, the domain expands from having a length scale of order the Planck Length to one of $\sim 10^{10^{12}}$. Thus, the domain scale has become much larger than the characteristic radius of the observable universe ~ 3000 Mpc $\approx 10^{28}$ cm (see § 8.3.2, Table 8.1 for the characteristic radius of the observable universe). The original small domain was in thermodynamic equilibrium (i.e., at the same temperature) and this is maintained during the inflation, except for small variation due to quantum fluctuations. The retention of near thermodynamic equilibrium solves the horizon problem; the small variations account the variations in the CBR temperature. As discussed in § 8.7.2, inflations flattens the domain, perhaps too much to be consistent with the observed Ω_0 and the age and formation time of the oldest stars. However, a relic of unexhausted scalar field energy might be tunable to provide a cosmological constant term which would make $\Omega_{0,\rm eff}$ nearly equal to 1. After inflation ends the ordinary big bang occurs and Friedmann evolution begins.

This inflationary scenario accounts for the observable universe with the only price being the invention of a scalar field with the right properties. The domain in which the observable universe is embedded is so immense that it is extremely unlikely that we are near enough to a domain wall to see beyond the domain. What is beyond the domain. I think that if one wishes nothing could be beyond. The domain is the entire universe and the initial domain was created *ex nihilo* just as one can suppose for the ordinary big bang picture. However, Linde sees the domain as embedded in indefinitely large, indefinitely long existing universe. If one imagines this universe as being infinite and eternal, then the need to specify all boundary conditions vanishes. Now Linde (Lin) is pretty vague, it seems to me, about his universe. Do the domains abut each other or are they separated by uninflated regions where the scalar field has high energy? Do expanding or inflating domains run into each other in any sense and if so what happens? Perhaps the space created in inflation and expansion is not to be imagined as expanding balloons pushing on each other; perhaps a nutshell could hold an inflated domain if not an infinite universe. Can an inflation start in a domain? Would it be dangerous if one started near us? Linde does not seem to address these issues. He describes his universe as fractal-like. Now I vaguely understand fractals: the coast of Norway looks the same on all size scales from Skagerrak to Liv Ullmann. But this description of self-reproducing inflation does not help me much.

In the other domains physical laws could be different. Not all physical laws or else nothing could be said about other domains. However, particle physicists think that low energy physical laws (those applying at less than the Planck Energy????) might be variable depending on how inflation was initiated. Thus, the masses of the particles and the values of fundamental constants could be different in other domains????. The number of spatial dimensions could be different too: more or less than 3????. In many domains, it is not likely that life such as ours could exist. Linde advises against crossing any domain boundaries.

8.7.4 Inflation: the Sad End of the Story?

If self-reproducing inflation or something like is correct, we are hopelessly embedded in our nutshell: i.e., our domain. The assumption of older cosmologies that the universe everywhere is like the one we observe is then invalid. Astronomy cannot alone lead to a TOE since we cannot see what the universe is really like. Since particle physics got us into this predicament, perhaps it can get us out. Maybe the next generation of high energy particle accelerators, or the one after that, or next still, *et cetera* will establish a more fundamental view. But there is room for doubt: to quote another inflation expert, Andreas Albrecht:

> Does it make any sense to discuss an absolutely fundamental TOE? I have suggested that even if one could construct such a theory, the observations we make are so superficial that

> the TOE might not constrain these observations in a very substantial way (Albrecht 1995, p 331).

I take it Albrecht means that the observable universe is "geography," not "physics."

It is possible that having wrested??? cosmology from the philosophers in the 17th century, science will now have to give it back? Such a tragic conclusion is I think premature. Albrecht again:

> The field of cosmology has a grand history of pushing back the boundary between physics and metaphysics. Decades ago, who would have thought that origin of relative abundances of the different chemical elements would be considered the subject of physical calculation rather than metaphysical speculation (Albrecht 1995, p. 331)!

To carry on this train: only in 1888??? was radio discovered: X-rays and radioactive decay in 1895???. And general relativity, the galactic nature of the spiral nebulae, the expansion of the universe, the cosmic background radiation, and inflation have all appeared in a human lifetime: Sinatra's lifetime to be specific. Not yet must we sink ourselves in a domain; the paths of space and time are open to our inquiring minds.

8.7 Life in the Univserse

Apppendix A: Esoterica

A1. The Secular Increase in the Mean Lunar Month

The tidal interaction of Earth and Moon cause slow secular increases the Earth's solar day by $\sim 2 \times 10^{-5}$ seconds per year and the lunar distance by ~ 3 centimeters per year (Fr193). The civil day is now defined to be exactly 86400 seconds and the second is determined by atomic clocks. To keep mean solar days and civil days synchronized extra seconds are added or subracted about once a year usually at midnight on December 31 or June 30 by the authority of the Bureau International de l'Heure in Paris (Ab126).

The secular increase in the lunar distances causes a secular increase in the mean lunar month, currently 29.530588 days (i.e., civil days). The relation between the two can be derived as follows. Kepler's third law relating the sidereal period, P, and semi-major axis (mean radius), a, of a body in orbit is

$$P \propto a^{3/2} . \tag{A1}$$

From calculus, small changes in P and a (i.e., ΔP and Δa) are related to first order by

$$\frac{\Delta P}{P} \approx \frac{3}{2} \frac{\Delta a}{a} . \tag{A2}$$

A synodic period $P_{\rm syn}$ (e.g., the lunar month) is given by

$$P_{\rm syn} = \frac{P P_\oplus}{P_\oplus - P} , \tag{A3}$$

where $P_\oplus$ is the Earth's year. Again using calculus, we find to first order that

$$\Delta P_{\rm syn} \approx \frac{P_{\rm syn}^2}{P} \left(\frac{3}{2}\right) \frac{\Delta a}{a} , \tag{A4}$$

where $\Delta P_{\rm syn}$ is the small change in the synodic period due to a small change in a.

Using current values for a and $P_{\rm syn}$, and $\Delta a = 3\,$cm, we find for the lunar month that

$$\Delta P_{\rm syn} \approx 4 \times 10^{-9} \quad \text{days} . \tag{A5}$$

Thus, the current increase in the lunar month is $\sim 4 \times 10^{-9}$ days per year.

A2. Ellipse and Eccentric Circle Orbits

Because ellipse and eccentric circle orbits play a large role in the history of astronomy it is worthwhile to present some of the formulae that are used to describe them.

The equation for an ellipse is given by

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1 , \tag{A6}$$

where a and b being the semi-major and semi-minor axes, respectively. The foci of an ellipse are located on the x-axis: the distance from a focus to the origin is labeled c. Their x coordinates of the foci are given by

$$\pm c = \pm\sqrt{a^2 - b^2} , \tag{A7}$$

where the lower case is for the left focus and the upper case for the right focus). The eccentricity of an ellipse is defined by

$$\epsilon = \frac{c}{a} . \tag{A8}$$

The formula for an ellipse in polar coordinates is given by

$$r_\pm = \frac{a\left(1-\epsilon^2\right)}{1 \pm \epsilon\cos\theta} \, , \tag{A9}$$

where the angle is measured from the positive x-axis and the radius measured from the left (lower case) or right (upper case) focus. The small eccentricity expansion of eqution (A9) is given by

$$r_\pm = a\left[1 \mp \epsilon\cos\theta - \epsilon^2\left(1-\cos^2\theta\right) + \ldots\right] \, . \tag{A10}$$

An ellipse satisfies the following formula (which can also be regarded as an ellipse defintion)

$$r_- + r_+ = 2a \, . \tag{A11}$$

The angle-averaged mean radius of an ellipse is given by

$$\langle r_\pm \rangle = a\sqrt{1-\epsilon^2} \, . \tag{A12}$$

The perifocus-apofocus mean radius (which is the usual mean radius cited for an ellipse????) is given by

$$\bar{r}_\pm = a \, . \tag{A13}$$

The formula for an ellipse in polar coordinates with origin at the center of the ellipse (i.e., the intersection point of the semi-major and semi-minor axes) is given by

$$r_{\rm origin} = \frac{a\sqrt{1-\epsilon^2}}{\sqrt{1-\epsilon^2\cos^2\theta}} \, . \tag{A14}$$

For eccentric circles

$$\epsilon = \frac{c}{a} \tag{A15}$$

is the eccentricity, where a is the circle radius and c is the the displacement along the x-axis of the origin from the circle center. The radius equation from the origin is given by

$$r_\pm = a\left[\sqrt{1-\epsilon^2\left(1-\cos^2\theta\right)} \mp \epsilon\cos\theta\right] \, , \tag{A16}$$

where the lower case is for a leftward displacement by c and the upper case for a rightward displacement by c. The small ϵ expansion of equation (A16) is given by

$$r_\pm = a\left[1 \mp \epsilon\cos\theta - \frac{\epsilon^2}{2}\left(1-\cos^2\theta\right) + \ldots\right] \, . \tag{A17}$$

Note that the ellipse and eccentric small ϵ radius equation expansions are identical to first order in small ϵ.

A3. The Fibonnaci Sequence

The most famous problem in Fibonnaci's *Liber Abaci* poses a problem that I have reformulated (from Bo281) as follows:

> Say you have have immortal rabbits. Each pair of rabbits (of different sex, but the same age) produce a pair of offspring (of different sex) every month starting from when the rabbits are two months old. At the start of the year, there is only one pair of rabbits just born. What is the population of rabbit pairs at the start of each month from the first month on?

The solution is the famous Fibonacci sequence: 1, 1, 2, 3, 5, 8, 13, 21, ... which is given by the Fibonacci recursion relation:

$$I_n = I_{n-1} + I_{n-2} \ , \tag{A18}$$

where the sequence starts from $n = 0$, and I_{-1} is defined to be 0 and I_0 to be 1. It is clear what is happening. At the start of each month you have the old rabbit pairs from the month before plus the new-born pairs who are equal in number to the number of reproductive pairs (i.e., the total number of pairs from the month earlier). Even if some mortality were thrown in, the situation is daunting as the Australians well know.

In general, there seems??? to be no explicit solution for the sequence members: i.e., one cannot obtain I_n for a given n without using the Fibonacci recursion relation. However, an approximate explicit solution exact in the limit of n going to infinity (i.e., an asymptotic solution) can be obtained. To be a bit more general—ah, the curse of generality—let us find the asymptotic solution for the recursion relation

$$I_n = F_{-1}I_{n-1} + F_{-2}I_{n-2} \ , \tag{A19}$$

where we allow the I_n to be real (or even complex) numbers, but F_{-1} and F_{-2} are real numbers greater than or equal to zero.

Let us assume the exponential solution given by

$$u_n = Cf^n \ , \tag{A20}$$

where C and f are parameters to be determined. (The exponential solution can easily be generalized to a continuous solution by changing the integer index variable n to a continuous variable.) If we substitute u_n for I_n in equation (A19), cancel the C parameter (the recursion relation is homogeneous), and rearrange, we find the quadratic equation

$$f^2 - fF_{-1} - F_{-2} = 0 \ . \tag{A21}$$

The solutions are

$$f = \frac{F_{-1} \pm \sqrt{F_{-1}^2 + 4F_{-2}}}{2} \ . \tag{A22}$$

For the cases that interest us, we consider only the positive solution (i.e., the "+'-case solution) in equation (A22). From equation (A21), it follows that

$$D_{-1} \equiv \frac{F_{-1}}{f} \qquad \text{and} D_{-2} \equiv \frac{F_{-2}}{f^2} \qquad (unitrm23)$$

sum to 1 and are both in the range $[0, 1]$.

Equation (A20) with f given by equation (A22) provides a exponential solution to the recursion relation. The parameter C, however, is free. Thus we have a family of exponential solutions. Given any one initial value for I_n, we can fit C and find a particular exponential solution. There are, however, two initial values for the recursion relation: I_{-1} and I_0. Unless these differ by a factor of f, then there is no exact exponential solution to the recursion relation.

We can, however, obtain an asymptotic solution as advertized. This solution is an exponential solution. We define

$$C_n = \frac{I_n}{f^n} \ . \qquad (unitrm24)$$

Now substitute $C_n f^n$ expressions into equation (A19), divide through by f^n, and make use of equation(A23) to obtain

$$C_n = C_{n-1}D_{-1} + C_{n-2}D_{-2} \ . \qquad (unitrm25)$$

It follows from the properties of D_{-1} and D_{-2} that

$$C_n \in [C_{n-1}, C_{n-2}] \ . \qquad (unitrm26)$$

It is clear that we can go on defining $C_n f^n$ solutions which bound the I_n solution ever more tightly from above and below. The function

$$|C_{n+1} - C_n| \tag{A27}$$

decreases strictly with n and reaches zero at $n = \infty$. Thus there is a limiting C_∞ which gives an asymptotic solution for I_n. The limiting C_∞ by

$$C_\infty = \lim_{n\to\infty} \frac{I_n}{f^n} \ . \tag{A28}$$

We have shown an asymtotic exists, but unfortunately equation (A28) does not give an explicit way to evaluate C_∞ and I know of no other way to obtain C_∞. For given initial conditions I_N, and I_{N+1}, I think one must use the recursion relation and determine an approximate C_∞ from

$$C_\infty \approx \frac{I_n}{f^n} \qquad \text{for} \qquad n >> N \ . \tag{A29}$$

However, with this approximate C_∞, the asymptotic solution (good to within a factor that is a small as one wishes to make it) is

$$I_n^{\text{asy}} = C_\infty f^n \ . \tag{A30}$$

In the case of the Fibonacci sequence, the acceptable value f and the approximate value for C_∞ are

$$f = \frac{1+\sqrt{5}}{2} \approx 1.6180339887499 \tag{A31}$$

and

$$C_\infty \approx 0.447213595499957 \ , \tag{A32}$$

respectively. In Table A1 we show the Fibonacci number, its asymptotic counterpart, and the ratio of the asymptotic counterpart to the Fibonacci number for n from 1 to 40. The asymptotic sequence is very good for say $n \geq 6$. Rounded-off to the nearest whole number the asymptotic sequence becomes exact.

Table A1. Fibonacci Sequence and Approximate Fibonacci Sequence

Order	Number	Asymptotic Number	Ratio
1	1	0.72361	0.72360679774998
2	1	1.17082	1.17082039324993
3	2	1.89443	0.94721359549996
4	3	3.06525	1.02174919474995
5	5	4.95967	0.99193495504995
6	8	8.02492	1.00311529493745
7	13	12.98460	0.99881516421149
8	21	21.00952	1.00045330924995
9	34	33.99412	0.99982695967642
10	55	55.00364	1.00006611133177
11	89	88.99775	0.99997475002523
12	144	144.00139	1.00000964496870
13	233	232.99914	0.99999631599888
14	377	377.00053	1.00000140717038
15	610	609.99967	0.99999946250979
16	987	987.00020	1.00000020530314
17	1597	1596.99987	0.99999992158120
18	2584	2584.00008	1.00000002995332

19	4181	4180.99995	0.99999998855885
20	6765	6765.00003	1.00000000437013
21	10946	10945.99998	0.99999999833076
22	17711	17711.00001	1.00000000063759
23	28657	28656.99999	0.99999999975646
24	46368	46368.00000	1.00000000009302
25	75025	75025.00000	0.99999999996447
26	121393	121393.00000	1.00000000001357
27	196418	196418.00000	0.99999999999481
28	317811	317811.00000	1.00000000000198
29	514229	514229.00000	0.99999999999924
30	832040	832040.00000	1.00000000000029
31	1346269	1346269.00000	0.99999999999989
32	2178309	2178309.00000	1.00000000000004
33	3524578	3524578.00000	0.99999999999998
34	5702887	5702887.00000	1.00000000000001
35	9227465	9227465.00000	1.00000000000000
36	14930352	14930352.00000	1.00000000000000
37	24157817	24157817.00000	1.00000000000000
38	39088169	39088169.00000	1.00000000000000
39	63245986	63245986.00000	1.00000000000000
40	102334155	102334155.00000	1.00000000000000

Note.—The first column gives the order of the Fibonacci number, the second, the Fibonacci number itself, the third, the asymptotic Fibonacci number, and the fourth, the ratio of the asymptotic counterpart to the Fibonacci number.

Appendix B: Constants and Formulae

Table B1. Numerical and physical constants

Quantity	Symbol	Value	Refs.
Base of Natural Logarithms	e	2.71828 . . .	
Inverse of Base e	e^{-1}	0.36788 . . .	
Pi	π	3.14159265 . . . $\approx$ 3.1416	
Speed of Light	c	$2.99792458 \times 10^{10}\,\mathrm{cm\,s^{-1}}$	Li1-1
Gravitational Constant	G	$6.67259(85) \times 10^{-8}\,\mathrm{cm^{-3}\,g^{-1}\,s^{-2}}$	Li1-1
1 megaton (explosive yield of TNT)		4.16×10^{22} ergs	WoA-20

Note.—The quantities in parentheses are the uncertainties in the last digits of the given values. The speed of light is exact by definition: one meter is defined to be the distance light in vacuum travels in 1/299792458 s.

Table B2. Time

1 (mean solar) day = 24 hr = 1440 min = 86400 s exactly	
1 sidereal day = 86164.0906 s	(Cl67)
1 mean lunar sidereal period= 27.321661 days	(Li14-4)
1 mean lunar or synodic month = 29.530588 days	(Mo731)
1 (mean) Hijra year (Hgr) = 12 mean lunar months = 354.36705	(derived from Gi46–47)
1 Egyptian year (Eyr) = 365 days exactly	(Ne81)
1 ideal year (iyr) = 365.2421875 days exactly	
1 (mean) tropical year (tyr) = 365.24219878 days = $1.0044881515\pi \times 10^7$ s $\approx \pi \times 10^7$ s	(Mo731)
1 (modern) mean Gregorian year (Gyr) = 365.24225 days exactly	
1 mean Omaric year (Oyr) = 365.242424 . . . days exactly	
1 original mean Gregorian year (OGyr) = 365.2425 days exactly	
1 mean Julian year (Jyr) = 365.25 days exactly	
1 (mean) sidereal year (syr) = 365.25636556 days	(Mo731)

Note.—In the Julian calendar the year is either 365 days (common year) or 366 days (leap year). Leap years are years whole number divisible by 4. The original Gregorian calendar is the same as the Julian calendar, except that century years not whole number divisible by 400 are common years, not leap years (Ab129–130). The modern Gregorian calendar is the same as the original Gregorian calendar, except that century years whole number divisible by 4000 are common years also (Ab129–130).

The Omaric calendar is like the Julian calendar, except that it omits a leap year every 132 years. The term Omaric calendar is my own. Omar Khayyam, Persian mathematician, astronomer, and possibly poet, (c. 1050–1123) was one of a committee who devised this calendar in 1074 (Joseph 1991, p. 309ff; Bo264ff). The ideal calendar (again my own name) is like the Julian calendar, except that it omits a leap year every 128 years. The ideal calendar obviously offers the best simple synchronization rule between days and tropical years. It would take 88,650 years for the ideal calendar to run one day ahead of the solar calendar.

Table B3. The Earth

Symbol: $\oplus$

Mean radius = 6371.0 km	(Li14-6)
Polar radius = 6357 km	(Cl68)
Equatorial radius = $R_{\oplus}^{\rm Eq}$ = 6378.140 km	(Li14-1)
Mass = 5.9742×10^{27} g	(Li14-6)
Mean density = 5.515 g cm^{-3}	(Li14-6)
Rotational velocity of the Earth at the equator = 0.46512 km	(Li14-6)
Mean lunar equatorial parallax=57′2.60″	(derived from Li14-1,-4)
Obliquity of the ecliptic = 23°26′21.448″ = 23.4392911°	(Li14-1)
General precession of the equinoxes rate = 1.3969712° per Julian century	(Li14-1)
General precession of the equinoxes period = 25770.037 Jyr	(derived from Li14-1)

Note.—The obliquity, the general precession rate, and the general precession period are based on the epoch 2000 values. The Jyr symbol means Julian years. Naturally, there may be secular and other variations that make the actual general precession period somewhat different and in fact variable.

Table B4. The Sun

Symbol: ⊙	
Radius = 6.9599×10^{10} cm	(Li14-2)
Mass = 1.9891×10^{33} g	(Li14-2)
Mean density = 1.409 g cm^{-3}	(Li14-2)
Luminosity = 3.86×10^{33} erg s^{-1}	(Li14-2)
Sidereal rotation period at 16° = 25.38 days	(Cl68)
Mean distance to the Sun = 1 Astronomical Unit = 1 AU = 1.4959787×10^{13} cm	(Li14-6)
Mean solar equatorial parallax=8.794148″	(Li14-1)

Table B5. The Moon

Symbol: Mo	
Radius = 1.738×10^{8} cm = = $0.2725 R_{\oplus}^{\rm Eq}$	(Li14-4)
Mass = 7.3483×10^{25} g = $0.01230002 M_{\oplus} = \frac{1}{81.30068} M_{\oplus}$	(Li14-4)
Mean density = 3.34 g cm^{-3}	(Li14-4)
Orbital period = 27.321661 days	(Li14-4)
Albedo = 0.14	(Li14-4)
Mean distance to the Moon = 3.844×10^{13} cm = $60.2684 R_{\oplus}^{\rm Eq}$	(Li14-4)
Mean lunar equatorial parallax=57′2.60″	(derived from Li14-1,-4)

Note.—The Moon's orbital period and day are the same length on average. Albedo is the fraction of light the Moon's surface reflects in the

Table B6. The planets I

Planet	$a_{\odot}$ (AU)	$P_{\rm sid}$ (days/Jyr)	ϵ (°)	θ_{ec} (°)	Albedo	No. of satellites
Mercury (Me)	0.38710	87.97 d	0.2056	7.004	0.106	0
Venus (Ve)	0.72333	224.70 d	0.0068	3.394	0.65	0
Earth (⊕)	1	365.26 d	0.0167	0	0.367	1
Mars (Ma)	1.52369	686.98 d	0.0933	1.850	0.150	2
[Ceres (Ce)]	2.7671	4.603 y	0.077	10.6		
Jupiter (Ju)	5.20283	11.86 y	0.048	1.308	0.52	16
Saturn (Sa)	9.53876	29.46 y	0.056	2.488	0.47	16
Uranus (Ur)	19.19139	84.07 y	0.046	0.774	0.51	18

Neptune (Ne)	30.06107	164.82 y	0.010	1.774	0.41	8
Pluto (Pl)	39.52940	248.6 y	0.248	17.15	0.3	1

Note.—The data in this table are a conflation of the data from Li14-3 and Ab693. The quantities $a_{\odot}$, $P_{\rm sid}$, ϵ, and θ_{ec} are mean orbital radius from the Sun (semimajor axis), sidereal period, eccentricity, and inclination of the orbit from the ecliptic, respectively. The symbol Jyr stands for Julian year. All the planets have traditional symbols, but except for that of Earth (e.g., $\oplus$), they cannot be reproduced with TEX. These symbols are little used, except again for the Earth symbol. Ceres is the largest asteroid (or minor planet) and the first discovered: we include it here for comparison.

Table B7. Largest asteroids

Asteroid	Discovery Year/No.		$a_{\odot}$ (AU)	ϵ (°)	θ_{ec} (km)	Diameter	Albedo	Class
Ceres	1801	1	2.77	0.08	11	940	0.10	C
Pallas	1802	2	2.77	0.23	35	540	0.14	Irr
Vesta	1807	4	2.36	0.09	7	510	0.38	Irr
Hygeia	1849	10	3.14	0.12	4	410	0.07	C
Interamnia	1910	704	3.06	0.15	17	310	0.06	C
Davida	1903	511	3.18	0.17	16	310	0.05	C
Cybele	1861	65	3.43	0.11	4	280	0.06	C
Europa (II)	1858	52	3.10	0.11	7	280	0.06	C
Sylvia	1866	87	3.48	0.09	11	275	0.04	C
Juno	1804	3	2.67	0.25	13	265	0.22	S
Psyche	1852	16	2.92	0.14	3	265	0.10	M
Patientia	1899	451	3.07	0.07	15	260	0.07	C
Euphrosyne	1854	31	3.15	0.23	26	250	0.07	C
Eumonia	1851	15	2.64	0.19	12	240	0.19	S
Bamberga	1892	324	2.69	0.34	11	235	0.06	C
Camilla	1868	107	3.49	0.07	10	230	0.06	C
Herculina	1904	532	2.77	0.17	16	230	0.16	S
Amphitrite	1854	29	2.55	0.07	6	225	0.16	S
Doris	1857	48	3.11	0.06	7	225	0.06	C

Note.—The data in this table are derived from Abe514. The quantities $a_{\odot}$, ϵ, and θ_{ec} are mean orbital radius from the Sun (semimajor axis), eccentricity, and inclination of the orbit from the ecliptic, respectively. The discovery number is just the order of discovery. We have appended (II) to the name of Europa to distinguish this asteroid from the Galilean satellite of Jupiter of the same name.

Ceres is the largest asteroid (or minor planet) and the first discovered. It was discovered by Giuseppe Piazzi (1746–1826) in Palermo, Sicily on January 1 1801 and named Ceres for the tutelary goddess of Sicily (Pa352; No317, 425–426). Ceres is also the goddess of the breakfast bowl. William Herschel coined the name none-too-appropriate generic name asteroid (No426) for astronomical bodies of which Ceres was the first. The more appropriate, but longwinded, name minor planet now has a certain vogue. Asteroids are quite small: Jupiter's Ganymede (the largest satellite in the solar system) has diameter of 5262 km, the Moon 3476 km, Pluto (the smallest planet) 2400 km (Abe511-512); but Ceres has diameter of only 940 km.

Asteroids were originally named for Greco-Roman goddesses, but these names were soon exhausted (Abe224ff). Soon other names were employed: e.g., those of pets, daughters, astronomers, summer student assistants. Currently, there are more than 6000 discovered asteroids (???). It is estimated that there are $\sim 10^5$ asteroids with diameters of $\gtrsim 1$ km.

The asteroids mainly lie in a belt between 2.2 and 3.3 AU: this is well confined between the orbital radii of Mars (~ 1.52) and Jupiter (~ 5.20). Some are much further out and in. Icarus comes to about 2/3 of Mercury's orbital radius from the Sun (No455). It has a very large eccentricity and also approaches the Earth. There are also about 50 Earth-approaching asteroids ranging up to 20 km in diameter, but most are only 1 or 2 km in diameter. One of these, Hermes, came as close as about two Moon distances ($\sim 120R_\oplus$) in 1937 (No455); another came to $\sim 27R_\oplus$ in 1991 (Ze510). It is now very credible that an asteroid (or other body) of a few kilometers in diameter impacted the Earth 65 million years ago (Ze510-511). There have been proposals to make a systematic search for Earth-approaching bodies to give us warning of any that threaten turn dinocidal. Even defensive measures with nuclear weapons have been debated (No456).

There is a simple spectroscopic classification scheme for asteroids (Abe226): C for carbon-rich, S for stony, and M for metal-rich. The C class are most numerous and have low reflectivity (i.e., low albedo): the carbon compounds on there surfaces are rather dark. The S asteroids appear to be rich in silicate minerals and have higher reflectivities. The M asteroids may be mostly metal: Psyche is the prime example of this class. Some asteroids do not fit well into this scheme (e.g., Pallas and Vesta): we have labeled them Irr for irregular in the table. Vesta is particularly interesting since it has a basaltic surface indicating a volcanic past. Volcanism in a small body is remarkable. Eucrite meteorites that are believed to come from Vesta suggest that there were lava flows 4.4 to 4.5 billion years ago: viz. from a time just after the solar system formation. Vesta is the brightest of the main-belt asteroids and is visible to the naked eye.

Table B8. Cycles and periods

Name	Years	Mean Lunar Months	days	12-month Years	13-month Years	Refs.
8-year cycle	8 tyr	$98.94613 \approx 99$	2921.9376	5	3	Pa51
19-year cycle	19 tyr	$234.9971 \approx 235$	6939.6018	12	7	Pa51, 108
25-year cycle	25 Eyr	$309.0016 \approx 309$	9125	...	...	Ne95
Saros period	$\sim$ 18 Gyr	$309.0016 \approx 309$	6585.322	...	...	Pa59

Note.—The symbols tyr, Eyr, and Gyr stand for tropical year, Egyptian year, and mean Gregorian year, respectively. To give them there full names, the 19- and 25-year cycles are called the 19-year Metonic and 25-year Egyptian cycle. The 8- and 19-year cycles are for the intercalation of lunar months. The 25-year cycle was not used for intercalation as the Egyptian year is invariant; it was used merely to predict mean lunar phenomena. The difference between predicted date and occurrence could be as much as 2 days???? due to variations in the lunar motion. The Saros period is an nearly exact repeat period for all lunar and solar eclipses.

Table B9. Trigonometric functions.

$$\text{opposite} \quad \text{over} \quad \text{hypotenuse} = \frac{y}{h} = \sin\theta \tag{B1}$$

$$\text{adjacent} \quad \text{over} \quad \text{hypotenuse} = \frac{x}{h} = \cos\theta \tag{B2}$$

$$\text{opposite} \quad \text{over} \quad \text{adjacent} = \frac{y}{x} = \tan\theta = \frac{\sin\theta}{\cos\theta} \tag{B3}$$

$$h^2 = x^2 + y^2 \tag{B4}$$

is the Pythagorean theorem.

Note.—The ratios of the sides of a right triangle (a triangle with a right or 90° angle) are given by the trigonometric functions where the argument is one of the non-right angles. The hypotenuse is the side opposite the right angle. The argument angle is between the hypotenuse and the adjacent (side). The opposite (side) is opposite the argument angle. Here the hypotenuse, adjacent, and opposite have lengths h, x, and y, respectively.

Table B10. Small x expansions.

The finite geometric series (which is valid for all x)

$$\frac{1-x^{m+1}}{1-x} = \sum_{n=0}^{m} x^n = 1 + x + x^2 + \ldots + x^m \tag{B5}$$

The infinite geometric series (which is valid for $|x| < 1$ only)

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + \ldots \tag{B6}$$

The trigonometric function expansions (which are valid for all x????)

$$\sin x = \sum_{n=1}^{\infty} (-1)^n \frac{x^{2n-1}}{(2n-1)!} = x \left(1 - \frac{1}{6}x^2 + \frac{1}{120}x^4 - \frac{1}{5040}x^6 + \ldots\right) \tag{B7}$$

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 - \frac{1}{720}x^6 + \ldots \tag{B8}$$

$$\tan x = x \left(1 + \frac{1}{3}x^2 + \frac{2}{15}x^4 + \frac{17}{315}x^6 + \ldots\right) \tag{B9}$$

Note.—The x variable in all these expressions is measured radians. Also any series can be truncated at any term providing the error in omitting the succeeding terms is deemed sufficiently small.

Appendix C: Chronology

Date	Astronomical, Scientific, Philosophical	Other
c. −4900		Small temple at Eridu on nearly virgin soil, the temple is clearly forerunner of later Mesopotamian temples (Llo41–43)
−4713	January 1 at 1200UT is Julian day 0	Midst of Neolithic age
−3200		Earliest pictographic script at Uruk, first writing in history, ancestor of Mesopotamian cuneiform (Llo36, 55; Po52)
c. −3000	Construction of Stonehenge I begins (Fr35)	
c. −2750		The Flood: the great flood of Mesopotamian tradition (Llo91–92)
c. −2600		Gilgamesh king of Uruk sometime in c. 2550–2650 BC (Llo92)
c. −2075	Construction of Stonehenge III including the sarsen stone ring and the trilithon horseshoe (Fr35)	
c. −600	Thales of Miletos (c. 625–545 BC) active	
−584	On May 28, an eclipse of the Sun supposedly predicted by Thales, the prediction story is almost certainly untrue (Ne142–143; Pa99)	Battle between Lydians and Persians interrupted by the eclipse
c. −490		Athenians defeat Persians at Marathon
c. −480		Athenians led by Themistocles defeat Persians commanded by King Xerxes in the naval battle of Salamas, effective start of Athenian hegemony in Aegean area
c. −450	Socrates (469?–399 BC) active	
−431		Peloponnesian War begins between Athens and Sparta
−404		Peloponnesian War ends with Athens defeat and Sparta becomes dominant in mainland Greece
−399	Execution of Socrates after judicial condemnation	

c. −387	Plato (427–347 BC) founds his school, the Academy, in Athens	
−336		Hellenistic period begins, Alexander the Great accedes to the kingship of Macedon, begins his career of conquest (Ba31)
−323		Death of Alexander the Great (Ba31)
−322	Death of Aristotle of Stagira	
c. −300	Euclid active in Alexandria	
c. −280	Foundation of Museum (Museion) of Alexandria by King Ptolemaios Soter???? (Pe328)	
	Apollonios, Archimedes, Aristarchos, Eratosthenes, Ktesibios active (c. 3rd century)	
−212	Probable death of Archimedes by a Roman soldier during the capture of Syracuse	Roman General Marcellus takes Syracuse
c.−200		Roman dominance in western Mediterranean, death of Hannibal (247–183? BC) the Carthegenian General
	Hipparchos of Nicaea (c. 190–120 BC) active (Pe346)	Roman power increases in eastern Mediterranean
c. −130		Roman civil wars begin
	Poseidonios of Rhodes (c. 135–50 BC) active (Pe381)	
c. −45	Julius Caesar imposes the Julian calendar which was designed by Sosigenes of Alexandria, an Aristotelean philosopher	
−44 Mar 15		Julius Caesar assassinated by a conspiracy on the Ides of March, fresh outbreak of Roman civil war
−30		Cleopatra and Mark Antony commit suicide, end of political Hellenistic period, Octavius Caesar sole Roman ruler
c. −30		Octavius assumes the name Augustus, establishes effective monarchy which is roughly stable for about two centuries
c. 100	Ptolemy (Claudios Ptolemaios) born	Last great bout of Roman conquest under Trajan (98–117)
c. 145	Ptolemy writes the *Almagest* (original title: *Mathematical Systematic*	

	Treatise)	
c. 175	Ptolemy dies	
c. 200		Beginning of century of Roman civil war and near collapse of the empire
c. 300	Pappos of Alexandria active, last great Alexandrian mathematician, writer of *Commentaries on the Almagest*	Restoration of Roman Stability by Diocletian and Constantine (288?–337)
330		Constantinople founded by Constantine on the site of ancient Byzantium, it becomes capitol of the eastern Roman empire (called the Byzantine empire by moderns????)
360	Theon of Alexandria active 360–370, writer of commentary on the *Almagest* and a revision of Ptolemy's *Handy Tables* (the only extant version)	
415	Death of Hypatia of Alexandria (daughter of Theon), commentator on Ptolemy	
476		Romulus Augustulus???, last western Roman emperor deposed by Odoacer (434?–493) first barbarian ruler of Italy
632		Death of Muhammad (570?–632), thereafter Islam spreads by conquest over the Arabian Peninsula, Middle East, Persia, North Africa, Spain
700		Sometime in the course of the next century printing invented in China (Ts1)
732		Charles Martel (690?–741), grandfather of Charlemagne, defeats Muslim invasion at battle of Tours???
800		Charlemagne (742–814) crowned emperor the Holy Roman Empire (Ba203) which later in Medieval times becomes a essentially a German empire of little coherence, sometime in the preceding century Carolingian minuscule invented from which modern lower-case letters derive (Gies77)
875	Thābit ibn Qurra of Harrān (c. 826–901) active mainly in Baghdad as translator from the Greek and writer of astronomical works	
900	al-Battānī (Albategnius) of Harrān (c. 850–929) active, compiler of a famous *zij* (i.e., collection of astronomical tables with canons (i.e., instructions)	
975	Ibn Yūnus (c. 940–1009) active, one of the foremost Arabian stronomers,	

	compiler of *Hakemite Tables*	
1000	Gerbert of Aurillac (c. 950–1003) active as Pope Sylvester II, one of the first intermediaries of Muslim science in Europe	
1258		Hūlāgū Khān, grandson of Genghis Khān sacks Baghdad, ending the Caliphate and becoming becoming ruler of the Mideast (Īlkhān or western khān)
1259	Nāsir al-Dīn al-Tūsī (1201–1274), Persian astronomer in the service of Hūlāgū Khān and inventor of the Tūsī couple used by Copernicus, becomes director of the Marāgha Observatory (No192ff, Pe370ff)	
1292		Before this year eyeglasses invented in Italy, only convex lenses for aiding farsightness available (Gies227)
1300		Before this year, the mechanical clock had been invented in Europe (Gies211)
1350	Ibn al-Shātir (1304–1376), Arabian astronomer, active in Damascus, anticipated many of Copernicus' results (e.g., discarding the equant)	Black Death in Europe, maybe a third of Europeans die by 1400
1400		Tamerlane (Timur Lang????) (1336?–1405), Mongol conqueror of most of western and southern Asia, active
1420	Uleg Beg (c. 1400–1449), grandson of Tamerlane and governor of Turkestan, founds observatory in Samarkand from which issued a new collection of tables and a star catalog	
1449	Uleg Beg, who became Mongol emperor in 1447, assassinated by his son	
1450		The European invention of movable type printing by Johann Gutenberg in Mainz (Gies241)
1453	The fall of Constantinople accelerates the spread of ancient Greek texts and learning in western Europe	Constantinople conquered by Ottoman Turks under Muhammad II (1430–1483), sultan of Turkey (1451–1483), end of Eastern Roman Empire
1464	Death of Nicolaus Cusanus, philosopher and cardinal, supporter of the theory of the rotation of the Earth and the infinity of the universe	
1492		Columbus (1446–1506) reaches the Americas (Ba239)
1500???		Concave lenses available for correcting myopia (Gies227)

1517		Martin Luther (1483–1546) begins the Protestant Reformation (Ba727)
1519???		Ferdinand Magellan's (1480?–1521) expedition begins its circumnavigation of the world
1543	Death of Nicolaus Copernicus (1473–1543) and publication of *De Revolutionibus Orbium Coelestium* (Nuremburg 1543, Basel 1566, Amsterdam, 1617)	
1545		Start of the Council of Trent (1545–1563), a landmark in the Counter Reformation of the Roman Catholic Church (Ba276)
1564	Galileo born (Ko358)	Shakespeare born (Ba1112)
1571	Kepler born on December 27 in Weil-der-Stadt in Swabia in Germany (Ca29)	
1582	Gregorian calendar introduced in Roman Catholic countries by Pope Gregory XIII (Pa220; Ab129)	
1600	On January 1 Johannes Kepler starts from Graz to Prague to meet Tycho Brahe (Ca99–100; Ko285) Giordano Bruno (a Copernican) burnt in Rome as a heretic, but not for his Copernicanism (Ko451; Pa224)	
1609	Kepler's *Astronomia Nova* published (Ca141) containing Kepler's 1st and 2nd laws (Ca132, 134)	
1610	In March, Galileo's *Sidereus Nuncius* published (No330)	
1616	On March 5, the Roman Catholic Church condems Copernicanism (Ko462–463)	Shakespeare dies (Ba1112)
1618	On May 15, Kepler discovers his On May 23, Thirty Year's War begins in Prague (Ca251) 3rd law (Ca286)	
1619	Kepler's *Harmonice Mundi* published (Ca264) containing Kepler's 3rd law (Ca286)	
1627	Kepler's *Tabulae Rudolphinae* (*Rudolphine* Tables) published (Ca324)	
1630	Kepler dies on November 15 noon in Regensburg (Ca358)	
1632	In February, Galileo's *Dialogue* is published.	
1633	Galileo's trial and condemnation to house arrest for life (Ko500)	
1636???	Galileo's *Two New Sciences* published (Ko502)	
1642	Galileo's dies (Ko502)	
1643	Newton born on January 4 (1642 December 25 on the Julian still in use in England) (Ab49–50)	
1687	Newton's *Principia* published (Ko516)	
1721	Halley introduces (but not discovers) the darkness paradox (No377)	

1727	Newton dies (No366)	
1729	Bradley announces the discovery of stellar aberration (No383)	
1781	Herschel discovers Uranus (No399-400)	
1789	Herschel completes 1.22 m (48 in) telescope (No400)	French Revolution begins
1801	On January 1, G. Piazzi in Palermo discovers the first asteroid Ceres (No425–426)	
1815		On June 18 Napoleon defeated at Waterloo (Ba1378)
1838	F.W. Bessel at Königsberg discovers stellar parallax (No415, 419)	
1839		Photography definitely in existence with invention due to J. N. Nièpce and L.J.M. Daguerre, independently, W.H. Fox Talbot and J. Herschel (No442; Ba304)
1840	J.W. Draper of New York does first astrophotograhy by taking daguerrotypes of the Moon (No443)	
1845	Lord Rosse in Ireland with his 1.83 m (72 in) telescope discovers that some nebulae have spiral structure (No437)	
1846	J.G. Galle at Berlin discovers Neptune using prediction of U.J.J. Leverrier; J.C. Couch had made a similar independent prediction in 1843 (No428–429)	
1850	In the following decade, R.B. Bunsen and G.R. Kirchhoff develop spectroscopy, already known, into a very useful analysis tool (No423–424)	
1859	Kirchhoff initiates the use of spectroscopy as a fundamental tool of astrophysics (No424)	
1905	Einstein invents special relativity (No511)	
1912	V.M. Slipher at Lowell Observatory in Flagstaff, Arizona measures the first Doppler shift of a spiral nebula (No522)	
1914		World War I begins
1915	Einstein publishes his theory of general relativity (No511)	
1916???	E.E. Barnard discovers Barnard's star; the the star with the largest proper motion (Fr306; 325)	
1917	Einstein publishes the first general relativistic model of the universe (No513, 520)	
1918		World War I ends
1919	J. Perrin speculates that nuclear fusion powers stars (No465)	Treaty of Versailles concluded (Ba1351)
1922	A.A. Friedmann publishes his first great paper on general relativistic cosmology (No524–525)	
1923	E. Hubble at Mt. Wilson Observatory proves that spiral nebulae are extragalactic star systems (i.e., other galaxies) (No509–510)	

c. 1926	Quantum mechanics invented by W. Heisenberg, E. Schrödinger, M. Born, et al.	
1929	E. Hubble discovers the expanding universe and Hubble's law (No523)	
1930	C.W. Tombaugh at Lowell Observatory in Flagstaff, Arizona discovers Pluto (No430–431)	
1932	K.G. Jansky initiates radio astronomy while working for Bell Telephone Laboratory in Holmdel, New Jersey (No545)	
1939	L. Meitner, O. Frisch, O. Hahn, and F. Strassman first publish on nuclear fission (Car441), H. Behte and C.F. von Weizäcker independently announce the carbon which turns out to be the the main nuclear fusion process in stars (No465, 531) (38 or 39???)	World War II begins
1943	C.K. Seyfert at Mt. Wilson discovers Seyfert galaxies (No558)	
1945		World War II ends, first use of nuclear bombs
1948	G. Gamow, R. Alpher, and R. Herman invent big bang cosmology and in 1949 Alpher and Herman predict the cosmic background radiation (No559; Ze484), F. Hoyle, H. Bondi, and T. Gold at work developing the steady-state universe model (No538)	
1957	On October 4, the Soviet Union launched Sputnik, the first artificial satellite to orbit the Earth, considered to be the dawn of the space age (No572)	
1961	On April 12, Y. Gagarin makes first manned space flight and orbit of the Earth (No576)	
1963	M. Schmidt, L. Greenstein, and T.A. Matthews discover the first quasar (No555–556)	
1965	A. Penzias and R. W. Wilson discover the cosmic background radiation (No560–561)	
1967	S.J. Bell, A. Hewish, et al. discover the first pulsar (No564–565)	
1969	On July 21, N.A. Armstrong and E.E. Aldrin of the U.S. make first manned landing on the Moon (No576)	
1978	J. Christy at the Naval Observatory in Flagstaff, Arizona discovers Pluto's moon, Charon (Ze242)	
1979	A.A. Starobinsky and A. Guth independently invent inflationary cosmology (Lin51; Ov245–247)	
1990	On April 24, the Hubble Space Telescope (HST) is put in orbit (Ze118)	
1992	the COBE satellite discovers fluctuation in the cosmic background radition (No613–614), A. Wolszczan and D.A. Frail discover the first pulsar	

	planets (Ze394)
1995	M. Mayor and D. Queloz discover the planet around a normal star (e.g., Naeye 1996)

Note.—For brevity, the BC dates in the date column are shown as negative numbers. If the year is n, then what is meant is that that year is the nth year counting from zero epoch. Consequently, n years from the zero epoch are not completed until the end of year n. Also consequently, centuries do not end until the end of a century year; thus the 21th century began???? at midnight January 1 2001. On the average the period between AD and BC dates is the former plus the later minus 1: e.g., the average period between 2 AD and 2 BC dates is $2 + 2 - 1 = 3$.

In these lecture notes we use BC and AD to designate whether the year is before or after the zero epoch of the conventional calendar. We omit the AD from dates which are clearly AD from context. The abbreviations BCE (before common era) and CE (common era) are sometimes used instead of BC and AD, respectively. However, BC and AD have a more general currency and there seems no point in obscuring the fact that the zero epoch of the calendar comes out of the Christian tradition: BC meaning before Christ and AD, anno Domini (the year of our Lord). This zero epoch was introduced by Dionysius Exiguus, an archivist in the Pope's service, in a report of 520 (viz. 520 AD) (Pa218). The zero epoch was supposed to be fixed by the year of the birth of Jesus, but Dionysius Exiguus was mistaken. It has not been possible to fix the time of that event. The use of AD dating only gradually became established. The English monk, the Venerable Bede of Jarrow (672–735), through his writings, was influential in spreading its use (No228).

References

Abell, G. O. 1982, *Exploration of the Universe* (Philadelphia: Saunders College Publishing) (Ab)

Abell, G. O., Morrison, D. & Wolff, S. C. 1988, *Realm of the Universe* (Philadelphia: Saunders College Publishing) (Abe)

Adamczewski, J., & Piszek, E. J. ????, *Nicolaus Copernicus and his Epoch* (Philadelphia: Copernicus Society of America) (Ada)

Adler, R., Bazin, M., & Schiffer, M. 1975, *Introduction to General Relativity* (New York: McGraw-Hill Book Company) (Ad)

Albrecht, A. 1995, "The Theory of Everything vs the Theory of Anything" in *Birth of the Universe and Fundamental Physics*, ed. F. Occhionero (Berlin: Springer-Verlag), 323

Barnhart, C. L., & Stein, J. (eds.) 1960, *The American College Dictionary* (New York: Random House) (Ba)

Boyer, C. B. 1968, *A History of Mathematics* (New York: John Wiley & Sons) (Bo)

Breszny 1996, "Breszny's Real Astrology" in *Nashville Scene*, January 11, 58

Cardwell, D. 1994, *The Norton History of Technology* (New York: W. W. Norton & Company) (Car)

Caspar, M. 1993 (1948), *Kepler,* (German First Edition 1948) tr. C. D. Hellman, Intro. O. Gingerich, Biblio. O. Gingerich, & A. Segonds (New York: Dover Publications, Inc.) (Ca)

Chambers, M. 1966, "The Crisis of the Third Century" in *The Transformation of the Roman World*, ed. L. White, Jr. (Berkeley: University of California Press), 30

Clark, J. B., Aitken, A. C., & Connor, R. D. 1962, *Physical and Mathematical Tables* (Edinburgh: Oliver and Boyd) (Cl)

Cohen, H. F. 1994, *The Scientific Revolution: A Historiographical Inquiry* (Chicago: University of Chicago Press) (Co)

Coles, P. & Lucchin, F. 1995, *Cosmology: The Origin and Evolution of Cosmic Structure* (Chichester: John Wiley & Sons) (Col)

Copernicus, N. 1543, *De Revolutionibus Orbium Coelestium* (Nuremburg: Johannes Petreius) (see also Ro)

DeVorkin, D. H. 1982, *The History of Modern Astronomy and Astrophysics: A Selected, Annotated Bibliography* (New York: Garland Publishing, Inc.)

Drake, S. 1978, *Galileo at Work: His Scientific Biography* (Chicago: University of Chicago Press) (Dr)

Dreyer, J. L. E. 1953 (1905) *A History of Astronomy from Thales to Kepler* (New York: Dover Publications, Inc.) (Dre)

Eisenstein, E. L. 1979, *The Printing Press as an Agent of Change: Communications and Cultureal Transformations in Early-Modern Europe* (2 vols.) (Cambridge: Cambridge University Press)

Flachière, R., & Chambry, E. (trs.) 1966, *Plutarque: Vies: Tome IV* (Paris: Société d'Edition << Les Belles Lettres >>)

Friedlander, M. W. 1985, *Astronomy from Stonehenge to Quasars* (Englewood Cliffs, New Jersey: Prentice-Hall, Inc.) (Fr)

Furley, D. 1987, *The Greek Cosmologists: Volume 1: The Formation of the Atomic Theory and its Earliest Critics* (Cambridge: Cambridge University Press) (Fu)

Gamow, G. 1970, *My World Line: An Informal Autobiography* (New York: The Viking Press) (Ga)

Gingerich, O. 1992, *The Great Copernicus Chase and Other Adventures in Astronomical History* (Cambridge, Massachusetts: Sky Publishing Corporation; Cambridge: Cambridge University Press) (Gi)

Gies, F., & Gies, J. 1994, *Cathedral, Forge, and Waterwheel: Technology and Invention in the Middle Ages* (New York: HarperCollins Publishers, Inc.) (Gies)

Hall, A. R. 1992, *Isaac Newton: Adventurer in Thought* (Oxford: Blackwell Publishers) (Ha)

Heath, T. L. 1991 (1932) *Greek Astronomy* (New York: Dover Publications, Inc.) (He)

Hecht, E., & Zajac, A. 1974, *Optics* (Reading, Massachusetts: Addison-Wesley Publishing Company) (Hec)

Hoffleit, D., Jaschek, C., & Saladyga, M. 1982, *The Brightest Star Catalogue* (New Haven, Connecticut: Yale University Observatory) (Ho)

Hoxie, A. 1966, "Mutations in Art" in *The Transformation of the Roman World*, ed. L. White, Jr. (Berkeley: University of California Press), 266

Jeffery, D. J. 2000, *Selected Wandering Poems* (Port Colborne, Ontario: Portpentagram Publishing) (Je)

Joseph, G. G. 1991, *The Crest of the Peacock: Non-European Roots of Mathematics* (London: I. B. Taurus & Co. Ltd Publishers)

Knuth, D. E. 1986, *The TEXbook* (Reading Massachusetts: Addison Wesley Publishing Company)

Koestler, A. 1977 (1959), *The Sleepwalkers: A History of Man's Changing Vision of the Universe* (Harmondswoth, Middlesex, England: Penguin Books, Ltd.) (Ko)

Kramer, S. N. 1967, *Cradle of Civilization* (New York: Time Inc.) (Kr)

Kuhn, T. S. 1970, (1957) *The Copernican Revolution: Planetary Astronomy in the Development of Western Thought* (Cambridge, Massachusetts: Harvard University Press) (Ku)

Kuhn, T. S. 1962, *The Structure of the Scientific Revolutions* (Chicago: The University of Chicago Press)

Lapidus, I. M. 1988, *A History of Islamic Societies* (Cambridge: Cambridge University Press) (La)

Levine, P. 1966, "The Continuity and Preservation of the Latin Tradition" in *The Transformation of the Roman World*, ed. L. White, Jr. (Berkeley: University of California Press), 206

Lide, D. R., & Frederikse, H. P. R. (eds.) 1994, *CRC Handbook of Chemistry and Physics* (Boca Rotan, Florida: CRC Press, Inc.) (Li)

Linde, A. 1994, "The Self-Reproducing Inflationary Universe," *Scientific American*, November, 48 (Lin)

Lloyd, S. 1984, *The Archaeology of Mesopotamia: From the Old Stone Age to the Persian Conquest* (London: Thames and Hudson Ltd.) (Llo)

Motz, L., & Duveen, A. 1977, *Essentials of Astronomy* (New York: Columbia University Press) (Mo)

Naeye, R. 1996, "2 New Solar Systems," *Astronomy*, April, 50

Neugebauer, O. 1969, *The Exact Sciences in Antiquity* (New York: Dover) (Ne)

North, J. 1994, *The Norton History of Astronomy and Cosmology* (New York: W. W. Norton & Company) (No)

Oppenheim, A. L. 1964, *Ancient Mesopotamia: Portrait of a Dead Civilization* (Chicago: University of Chicago Press) (Op)

Overbye, D. 1992, *Lonely Hearts of the Cosmos: The Scientific Quest for the Secret of the Universe* (New York: HarperPerennial) (Ov)

Pannekoek, A. 1961 (1951), *A History of Astronomy* (New York: Interscience Publishers, Inc.) (Pa)

Pedersen, O., & Pihl, M. 1974, *Early Physics and Astronomy* (Cambridge: Cambridge University Press) (Pe)

Postgate, J. N. 1992, *Early Mesopotamia: Society and Economy at the Dawn of History* (London: Routledge, Chapman, and Hall, Inc.) (Po)

Ptolemy (Klaudios Ptolemaios [Κλαύδιος Πτολεμαῖος]) c. 150 AD, *Almagest* (see also To)

Ronan, C. A., & Needham, J. 1981, *The Shorter Science and Civilization in China: Volume 2* (Cambridge: The Cambridge University Press) (Ron)

Rosen, E. 1978, Nicolas Copernicus: On the Revolutions (Baltimore, Maryland: The Johns Hopkins University Press) (Ro)

Rosen, E., & Hilfstein, E. 1995, *Copernicus and his Successors* (London: The Hambledon Press) (Ros)

Runciman, S. 1955, *A History of the Crusades: Volume III: The Kingdom of Acre* (Cambridge: The Cambridge University Press) (Ru)

Saliba, G. 1994, *A History of Arabic Astronomy: Planetary Theories during the Golden Age of Islam* (New York: New York University Press) (Sa)

Sayili, A. 1960, "Appendix II: The Causes of the Decline of Scientific Work in Islam" in *The Observatory and its Place in the General History of the Observatory* (Ankara)

Seeds, M. A. 1988, *Foundations of Astronomy* (Belmont, California: Wadsworth Publishing Company) (Se)

Shakespeare, W. 1979, *The Complete Illustrated Shakespeare,* ed. H. Staunton (New York: Park Lane)

Shapiro, S. L., & Teukolsky, S. A. 1983, *Black Holes, White Dwarfs, and Neutron Stars: The Physics of Compact Objects* (New York: John Wiley & Sons)

Shu, F. H. 1982, *The Physical Universe: An Introduction to Astronomy* (Mill Valley, California: University Science Books) (Shu)

Smith, E. v.P. & Jacobs, K. C. 1973, *Introductory Astronomy and Astrophysics* (Philadelphia: W. B. Saunders Company) (Sm)

Thoren, V. E. 1990, *The Lord of Uraniborg: A Biography of Tycho Brahe* (Cambridge: Cambridge University Press) (Th)

Toomer, G. J. (tr. and ed.) 1984, *Ptolemy's Almagest* (New York: Springer-Verlag) (To)

Toomer, G. J. 1996, "Hypatia of Alexandria," Journal of the History of Astronomy, May, 174????

Tsien Tsuen-Hsuin 1985, *Science and Civilization in China: Volume 5, Part I: Paper and Printing* (Cambridge: Cambridge University Press) (Ts)

Vryonis, S., Jr. 1966, "Hellas Resurgent" in *The Transformation of the Roman World*, ed. L. White, Jr. (Berkeley: University of California Press), 92

Wagner, J. K. 1991, *Introduction to the Solar System* (Philadelphia: Saunders College Publishing)

Wall, B. E. 1975, "Anatomy of a Precursor: The Historiography of Aristarchos of Samos" in *Studies in History and Philosophy of Science*, 6, 201 (Wa)

White, L., Jr. 1978, *Medieval Religion and Technology: Collected Essays* (Berkeley: University of California Press)

Wolfson, R., & Pasachoff, J. M. 1990, *Physics Extended with Modern Physics* (Glenview, Illinois: Scott, Foresman and Company) (Wo)

Zeilik, M. 1994, *Astronomy: The Evolving Universe* (New York: Wiley) (Ze)

Index

Note.—Items are referenced to sections, not pages. This convenient for updating the lectures.

Personal names have had to treated nonuniformly. Modern persons (i.e., people from circa 1500 on) have been listed by their surnames. Pre-modern persons and some modern persons have been listed by their most conventional name or abbreviated name. Thus Aristarchos of Samos (c. 3rd century BC) is listed under the a's and Nasīr al-Dīn al-Tūsī (13th century) is listed under the t's for Tūsı. Modern people who are listed by conventional names are Galileo (full name Galileo Galilei) and Tycho Brahe (who is usually just called Tycho).