Theme 3: Early Astronomy: Earth and Sky

3.1 Megalithic Astronomy

Britain and neighbouring areas of northern Europe boast a large number of Neolithic stone monuments, the most spectacular and well-known being Stonehenge. These monuments are generally accepted as having astronomically significant alignments, though the number and accuracy of these alignments are much disputed. There is no doubt that Stonehenge is aligned on the midsummer sunrise–midwinter sunset line, and two impressive chambered tombs (Newgrange in Ireland and Maeshowe on Orkney) are aligned so that the midwinter sunrise (at Newgrange) or sunset (at Maeshowe) illuminates the main chamber. Whether these alignments are "astronomical" or "ritual" is a different question: most Christian churches are aligned east-west, which could be associated with the equinoxes, but this does not mean that Christian astronomer-priests use them to make detailed astronomical observations...

Lunar alignments are also reasonably convincingly demonstrated at Stonehenge, as shown in figure 3.1, and a number of Scottish megalithic sites, including the impressive set of standing stones at Callanish on the Isle of Lewis, are claimed to incorporate lunar alignments, though the precision of these is much debated. It appears that the megalithic culture was interested in the northern and southern limits of moonrise and moonset, which exceed the solar limits because the Moon's orbit is inclined to the ecliptic.

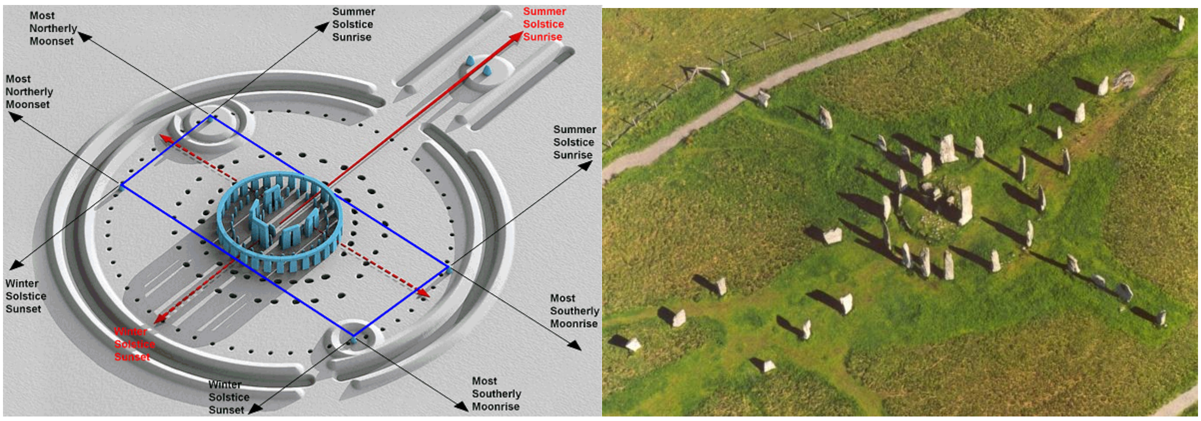

Figure 3.1: Left, Stonehenge and its principal alignments; right, an aerial view of the main structure at Callanish (as with Stonehenge, this is part of a "Neolithic landscape" including several subsidiary circles). Figures from http://calgary.rasc.ca/calgary-stonehenge.htm and https://www2.stetson.edu/neolithic-studies/stone-circles/callanish-from-the-air-isle-of-lewis-scotland/.

These alignments certainly could have been used for calendric purposes, although ritual use seems at least as likely. Thurston argues that "we should not glibly assume that Stonehenge was a temple", but the amount of communal effort that went into building this monument—including transporting the bluestones from Wales—seems rather excessive for a simple observatory! (It is also clear that a simpler and cheaper wooden structure would be *better* for astronomical purposes: alignments defined by enormous, difficult-to-erect standing stones are unlikely to be as precise as those defined by straightforward wooden posts, just on the grounds that making sure your stone ends up exactly where you intended it, within centimetres, is highly non-trivial with a Neolithic toolkit!) The appearance of both solar and lunar alignments would suggest that any calendric use involved a lunisolar calendar (see below); an interest in eclipses has also been suggested.

In conclusion, the alignments of megalithic monuments are sufficiently suggestive that we can safely deduce a keen interest in at least some aspects of the seasonal variations in the positions of sunrise, sunset, moonrise and moonset. Whether the concern was purely ritual or partly utilitarian (e.g. calendar-fixing) is unclear, as is whether there was any element of eclipse prediction (Fred Hoyle showed that a modern astronomer could devise a method of using Stonehenge for eclipse prediction: this is very different from demonstrating that it was designed to be so used!). The basic problem is that in the absence of written records it is extremely difficult to deduce the use of a complex structure purely from its architecture (it doesn't help that Stonehenge, by far the most extensively studied "astronomical" structure, is unique, and therefore conclusions cannot be tested by predicting the layout of Stonehenge analogues at other latitudes).

Fortunately, many other ancient societies with an interest in astronomical phenomena were literate, and their scripts can be read (the scripts we can't read are generally those where the available material is limited, e.g. Etruscan and the Indus valley scripts; by definition, these would not provide a great deal of information on those cultures' astronomy even if we could read them, since there isn't much to read!). We are therefore on much more secure ground in working out what these societies were doing, and why they were doing it.

3.2 The Calendar

The earliest arguably astronomical records are various series of marks, e.g. at the famous caves of Lascaux, which are interpreted as representing a lunar month. This demonstrates the importance of the night sky for calendrical purposes from earliest prehistory.

The important time periods for early societies are

- the day (and its seasonal variation in length, especially at relatively high latitudes, e.g. Britain);
- the year (the most relevant period for food production, whether by farming or by foraging);
- the lunar month (the most convenient time-keeper for short-term planning, especially since the changing phase of the Moon allows the time of month to be read off with fair accuracy).

Unfortunately, these quantities are not simply related:

- one tropical year (equinox to equinox) = 365.2422 solar days (slightly more in ancient times because the day was slightly shorter);
- one lunar synodic month = 29.5306 days, on average (it varies from 29.27 to 29.83, because of the eccentricity of the Moon's orbit);
- therefore one tropical year = 12.368 lunar months.

To make matters worse,

- the speed of the Sun around the ecliptic varies over the course of a year, because the Earth's orbit is not precisely circular;

- the length of the solar day also varies slightly through the year, for the same reason (Mean Solar Time corrects for this effect using the **Equation of Time**[1]);
- the **synodic month** (New Moon to New Moon) is not the only relevant lunar month[2]:
  - the **anomalistic month** (lunar perigee to lunar perigee), which relates to the speed with which the Moon moves, is 27.5545 days;
  - the **nodal**, **draconic** or **draconitic** month (successive returns of the Moon to the ascending node of its orbit relative to the plane of the ecliptic—the Moon's orbit is inclined at 5° to the ecliptic), which is relevant to the prediction of eclipses, is 27.2122 days;
  - in fact, the only lunar period which is *not* of great interest to early astronomers is the one we would regard as its "true" orbital period—the **sidereal month** of 27.3217 days!

Much of early astronomy is essentially concerned with the task of attempting to record, systematise and predict the effects of these various periods. For calendric purposes, the main aim of this work is to develop a set of concordances based on common multiples of the various quantities, so that the calendar can be codified and future dates accurately predicted. The observations that are actually recorded are in many cases more detailed than this, reflecting additional requirements (e.g. an astrological need to know exactly where the Moon is in the sky, or a desire to predict eclipses).

3.2.1 Eclipse cycles

Eclipses of the Sun and Moon must have been deeply alarming to ancient peoples, especially those for whom the Sun and Moon had divine status. Literate societies therefore tended to record such events: we have records of lunar eclipses kept by the Babylonians (and listed in the work of Ptolemy) from about 720 BC, and fairly continuous Chinese records from a similar date.

Using such records, it would eventually have become clear that eclipses occur in predictable cycles: this information was of great use to the ancient astronomers, both those working with the calendar and those concerned with astrology, and is still of use today (ancient eclipse records can be very useful in providing absolute dates to calibrate local calendars, and we still label solar eclipses according to which saros cycle they belong to). As a lunar eclipse can only occur at Full Moon when the Moon is near one of its nodes, it can be seen that the relevant cycles are common multiples of the synodic month and half the nodal month (we don't care which node we're at). The most commonly used cycle in ancient times was the **saros** of 223 synodic months (= 242 nodal months). By good fortune, the saros is also a good cycle for the anomalistic month (223 synodic months very nearly equal 239 anomalistic months), which helps to ensure that the chance of a repeat eclipse is very high. Both the Babylonians and the Greeks used the saros for eclipse prediction; the Chinese used a somewhat shorter cycle, 135 synodic months = 146½ nodal months (which is not quite as good: the difference between 223 synodic months and 242 nodal months is 0.0286 days, while that between 135 synodic months and 146½ nodal months is 0.0437 days; also, the Chinese cycle is not a good cycle for the anomalistic month).

Solar eclipses are much harder to predict, because of the extremely small size of the Moon's shadow on the Earth; though the saros is a good cycle, there is no guarantee that the expected solar eclipse will be visible from your city. This probably made it more difficult to recognise that solar eclipses were predictable: Needham, in *Science and Civilisation in China,* notes that the

---

[1] See https://en.wikipedia.org/wiki/Equation_of_time.
[2] See http://eclipse.gsfc.nasa.gov/SEhelp/moonorbit.html for a nice discussion of the various months, with plots and diagrams.

Shih Ching, which dates to somewhere around 800 BC and contains some astronomical material, refers to solar eclipses as "abnormal" and lunar as "usual"; he comments that this "suggests that the latter were expected but the former were not." On the other hand, predicting too many eclipses was probably a minor difficulty: as eclipses seem in general to have been bad omens (not unexpectedly—having the Sun suddenly disappear, or the Moon turn blood red, does not lend itself to optimistic interpretation), the professional disaster is to fail to predict one that does happen (if you predict one that doesn't happen, you explain that your propitiation rituals have worked!).

From a calendric as opposed to an astrological viewpoint, eclipses have the advantage that the central moment of the eclipse fixes precisely the instant of Full Moon (lunar eclipse) or New Moon (solar eclipse). Recording the precise times of eclipses in successive saros cycles therefore allowed ancient astronomers to refine their value of the length of the month.

Astronomically based calendar systems can be divided into three general categories: **lunar calendars** based on the Moon, **lunisolar calendars** based on the Moon but adjusted to take account of the solar year, and **solar calendars** based purely on the year. Our own Gregorian calendar is a solar calendar which descends from a lunisolar predecessor; it has kept the name "month" for its principal subdivisions, but lost any connection with the phases of the Moon.

3.2.2 Lunar calendars

A lunar calendar is in principle extremely straightforward:

- the month starts at a particular phase of the Moon (usually, but not invariably, the New Moon);
- the year has 12 lunar months (and thus lasts for 354.37 days).

The lunar phase may be calculated from a knowledge of the length of the synodic month, or directly observed. In the case of a directly observed definition, this calendar is extremely simple to implement, but practically impossible to codify (since the visibility of the first thin crescent of the Moon depends critically on the weather, the local geography, and the designated observer).

To codify a lunar calendar so that dates can be specified in advance, the usual starting point is to alternate 29 and 30-day months, giving a mean month length of 29.5 days. This is a little too short, and so an extra day will have to be inserted approximately every 30 months. For example, a system which is sometimes used to provide a calculated approximation to the Islamic lunar calendar inserts 11 such "leap days" in 30 lunar years: i.e. it amounts to saying that $30 \times 12 = 360$ synodic months is equal to $360 \times 29.5 + 11 = 10631$ days, or one synodic month = 29.5305555… days. This is correct to 4 decimal places.

Purely lunar calendars have one enormous drawback: the "lunar year" is just under 11 days shorter than the solar year, and the lunar calendar therefore quickly gets out of step with the seasons. This is a disaster for an agrarian society, where calendar-derived fixed points such as the date when your taxes are due *must* be reasonably well synchronised to the season (paying taxes in money is a relatively recent innovation: for most of history taxes were paid in kind, as a proportion of the harvest, a set number of cows, etc.). For this reason, essentially all ancient lunar calendars are really lunisolar calendars (see below). The one extant purely lunar calendar is actually derived from a lunisolar calendar, the intercalation process having been removed for religious reasons.

This lunar calendar is the Islamic calendar. It is an observation-based lunar calendar whose months start with the first observation of the crescent Moon after astronomical New Moon. Because it is observation-based, it does not have (or need) a formal rule for inserting "leap days" —instead, the lengths of months will vary rather unsystematically between 29 and 30 days (assuming uniformly good weather; some 31 or 28 day months will happen in the real world when the atmospheric conditions delay the sighting of the crescent Moon). It is entirely possible for calendars in different parts of the Muslim world to be out of step by a day or so, as a result of different weather conditions.

3.2.3 Lunisolar calendars

The lack of synchronisation between the lunar "year" of 354.37 days and the solar year of 365.2422 days can be solved by allowing "lunar leap years" of 13 months approximately every third year. This is a lunisolar calendar. Almost all astronomically motivated ancient calendars are of this type, e.g. the Babylonian, Jewish and Chinese calendars (there is also an Indian lunisolar calendar which works on somewhat different principles). What distinguishes different lunisolar calendars is the rule employed for deciding whether or not the current year has a 13th month. There are three qualitatively distinct categories of rule:

- *royal decree*: insert an extra month when the king says so (this tends to be unsatisfactory, since the king may have more immediate concerns, e.g. a war with the neighbours);
- *observational*: use a solar-year marker, e.g. the heliacal rising of a star (see below), and apply a rule of the form "if this event occurs close to the beginning of the appropriate lunar month, the year has 12 months; if it occurs near the end, it has 13 months";
- *systematic*: determine the true lengths of the lunar month and the solar year as accurately as possible, and find a good common multiple—then use this to design a pattern of intercalations.

From the point of view of advancing astronomy, the third system is the most productive, because the task of determining the mean length of the synodic month requires careful observation, good record-keeping and sophisticated mathematics. The reason for this is that the Moon's apparent speed across the sky varies according to the anomalistic month, not the synodic month: therefore the actual length of the synodic month varies significantly as noted above. Furthermore, the absolute time of astronomical New Moon is not directly verifiable by observation (except in the event of a solar eclipse), as the Moon is then invisible, and the time of Full Moon is difficult to judge with any accuracy (to the eye, the Moon remains full for some time). And finally, the length of the solar day is itself varying, because of the eccentricity of the Earth's orbit (this does not, of course, affect the Earth's rotation, but it does make a slight difference to the length of the solar day).

3.3 Babylonian Astronomy

Among ancient civilisations, the Babylonians had a significant advantage in astronomical calculations: a well-designed and effective system of numerical notation. They used a place-value system, similar to the one we have acquired from India via the Arabs, but based on 60 rather than 10. The reason for the base of 60, which the Babylonians inherited from the Assyrians, is unclear, but the actual notation used by the Babylonians makes it clear that there was an underlying base of 10 (for example, 23 would be written ‹‹|||, while 76 would be | ‹|||||| [1×60 + 16×1]). I suspect, in a system which did not have a notation for fractions, 60 was chosen because most simple fractions can be expressed in 60ths—e.g. ⅙ = 10/60, ⅕ = 12/60, ¼ = 15/60, ⅓ = 20/60, ½ = 30/60; another possibility is that it was influenced by the well-known geometrical construction which allows you to divide a circle into 6 equal parts. Regardless of its origins, this system lends itself to arithmetical calculations (in contrast to, say, the Roman system where 23 is XXIII, 76 is LXXVI, and their sum is XCIX). We can see the Babylonian system surviving in our degrees, minutes and seconds (which are exactly equivalent to theirs), and also in our hours, minutes and seconds (although the Babylonians actually divided the day into 360 "degrees" (same Babylonian word, *ush*, as their angular degree), or into 12 two-hour units called *beru*).

The Babylonians kept careful records of astronomical observations and calculations, and many of their records (written in cuneiform on clay tablets) have survived. It has been possible to work out that the Babylonians had a sophisticated numerical understanding of the motions of the Sun, Moon and naked-eye planets. They knew that the Sun's speed across the sky was not constant, and approximated this in two distinct ways: in "System A", the Sun travels at a constant speed for half the year, and a different constant speed for the other half; in "System B", the speed is represented by a zigzag, increasing uniformly to a maximum and then decreasing symmetrically to a minimum.

System B leads to simpler calculations and more accurate answers, leading to the suspicion that it was invented later (if it were the earlier system, why would anyone have bothered inventing System A?), but both systems were in use during the period for which we have the best records (the Seleucid era, from Alexander to the Parthians). The Babylonians also understood and kept records of eclipses of the Sun and Moon; these can be used to determine the average synodic month, since a lunar eclipse accurately determines the time of full Moon, and a solar eclipse, even if only partial, determines new Moon. Unlike the Greeks, however, the Babylonians seem to have been content with empirical calculational ansãtze: they did not attempt to construct models of the solar system which would explain the physical basis of the calculations.

The effect of these calculations was to convert an astronomically adjusted lunisolar calendar into a codified, predictable system. For astronomical adjustment, Pannekoek quotes an Assyrian text: "When at the first day of Nisannu the moon and star Mulmul [the Pleiades] stand together, the year is common; when at the third day of Nisannu the moon and star Mulmul stand together, the year is full." This suggests that the Pleiades are being used to decide whether a year should be 12 lunar months ("common") or 13 ("full"). The codified system, in contrast, uses the **Metonic cycle**—the fact that 19 solar years are almost exactly equal to 235 lunar months. Since 235 = 19×12 + 7, this means that 7 extra months must be intercalated into every 19 lunar years, in any suitable pattern (the Babylonians put them in years 3, 6, 8, 11, 14, 17 and 19). The Metonic cycle is named after a Greek astronomer Meton, who lived in the 5th century BC and suggested using it to regularise Greek calendars; it is not known whether Meton and the Babylonians discovered the cycle independently or not. The Babylonian use of the cycle is only attested in the Hellenistic period, when it could have been borrowed from the Greeks; if it was actually used earlier, there was sufficient contact between Greece and the Near East for the knowledge to have spread in either direction. Babylonian astronomy certainly had other influences on Greek astronomy: many of our modern constellations derive directly—even to their names—from Babylonian originals via their adoption by the Greeks, and Hipparchos seems to have made extensive use of Babylonian records (some of the numbers Ptolemy attributes to Hipparchos appear to have been calculated using Babylonian System B).

3.4 Egyptian calendar

Contrary to the beliefs of some New Age authors, the Egyptians were disappointingly unproductive from the point of view of ancient astronomy. This may be partly because their numerical system did not lend itself to complicated mathematics as the Babylonian sexagesimal system did, but is probably mostly because they appear to have been (somewhat like the Romans) rather uninterested in abstract knowledge. They were, however, quite good at calendars, because of the need to keep track of the Nile floods. From ancient times, the date when Sirius was first visible, rising just before sunrise (the **heliacal rising** of Sirius), was used as the astronomical herald of the Nile floods, and it was therefore the natural astronomical marker for the Egyptian lunisolar calendar. The rule is straightforward: Sirius is always supposed to rise in the twelfth month. Since the lunar "year" is 11 days shorter than the solar year, if in a given year Sirius rises in the last eleven days of the twelfth month, a thirteenth month must be added (otherwise, the next heliacal rising of Sirius will take place not in the twelfth month of the next year but the first month of the year after that).

It is testament to the pragmatic nature of the Egyptians that even this simple rule was obviously too inconvenient to appeal to the Egyptian civil service. As early as 3000 BC or so, the lunisolar calendar was therefore replaced for administrative purposes by a simple approximation to a pure solar calendar: 12 months of 30 days plus a 5-day end-of-year festival. There were no leap years, so this calendar gradually gets out of step with the seasons, but in the minds of Egyptian civil servants its convenience clearly outweighed this disadvantage.

Although the Egyptian civil calendar has no real astronomical basis other than approximating the solar year, it was widely used by ancient astronomers precisely because of its lack of intercalation (much as the Julian date is used by modern astronomers). For example, Hipparchos converted Babylonian lunar eclipse observations from the Babylonian system (which was somewhat chaotic in the earlier centuries because of adherence to the "royal decree" system of intercalation) to Egyptian dates for calculational convenience.

3.5 Chinese astronomy

The Chinese, in contrast to the Egyptians, have a long and distinguished astronomical tradition going back as far as that of Babylon. In China, astronomy was a branch of the civil service—the Emperor was viewed as possessing "the Mandate of Heaven", so clearly the heavens had to be carefully observed to ensure that any celestial portents were properly recognised and interpreted—and there was a position analogous to the original remit of the Astronomer Royal, in charge of the Imperial Observatory. An unfortunate side-effect of this is that Chinese astronomical records tended to be stored in the imperial palace and/or imperial government buildings: changes of dynasty in China were often accomplished by force, with associated sacking and burning of the said palaces, which means that many astronomical records did not survive. We are therefore often reliant on secondary sources, and dating of early Chinese astronomical knowledge is not very secure.

The Chinese recorded solar and lunar eclipses: there are occasional surviving records from before 1000 BC, and fairly continuous sets from 720 BC onwards. In modern times, Chinese historical records are mostly used to investigate transient phenomena such as comets, novae and supernovae, which were carefully if sometimes ambiguously recorded (it is not always obvious to the modern scholar if a given "guest star" is to be interpreted as a comet or a supernova!). They also had, at least from the first century BC onward, decent measurements of the synodic periods of the naked-eye planets, and an excellent value for the synodic month. While their models of the physics were much less sophisticated than those of the Greeks, their observational techniques were clearly excellent.

The Chinese lunisolar calendar started months at the new moon, and originally started the year at the winter solstice, though the civil calendar was changed around 104 BC to start in our February. They used the Metonic cycle to organise intercalated months: in a given 19-year cycle, years 2, 5, 8, 10, 13, 16 and 19 are 13-month years. This is very similar to the Babylonian system and may not be independent: indirect trade between China and the Near East via the Silk Route has a very long history[3].

3.6 The Indian lunisolar calendar

Indian astronomy has a very respectable history, both observational and theoretical. Alexander's empire reached India, and Indian theoretical astronomy was influenced by Hellenistic work—for example, Āryabhaṭa (born AD 476) developed a geocentric model of planetary motion, based on pre-Ptolemaic Greek models, but with innovations of his own: his epicycles were not of fixed size, and he explained the overall rotation of the night sky as resulting from the rotation of the Earth. The Indians also made essential mathematical advances, including the introduction of the number zero and our modern numeric system (the so-called Arabic numerals were adopted by Islamic scholars but came originally from India).

The description of the Indian lunisolar calendar I have found (webexhibits.org/calendars) dates only from 1957, and neglects a host of regional variations. The underlying astronomy is, however, quite within the capabilities of early Indian astronomers.

The calendar is based on a series of 12 solar "months", which are defined by the time taken for the Sun to traverse 30° of the ecliptic. Their lengths are therefore not whole numbers of days and are shorter near January, when the Earth is near perihelion. If this were the complete picture, the calendar would be purely solar. However, these solar months are not used in practice: instead, lunar months starting with the New Moon are used. The lunar month takes the name of the solar month in which its initial New Moon falls; an intercalation occurs automatically if a solar month contains two New Moons. (It very occasionally happens that one of the "short" solar months contains no New Moon: in such a case, that solar month's name does not occur in the year in question. This does not mean that the year will contain only 11 lunar months; it turns out that when this happens a preceding "long" solar month will have had two New Moons, and the year has twelve lunar months as usual.)

This is an unusual lunisolar calendar in that the solar aspect is actually the primary determinant, with the lunar months defined secondarily. From the Wikipedia entry on the Hindu calendar, it appears that this is how it was originally set up, although there are many regional variants concerning whether the month starts at New Moon or Full Moon and when the New Year is celebrated.

3.7 Solar Calendars

The most natural calendar for an agrarian society is a pure solar calendar. However, in preliterate societies this is difficult to keep track of: there is no real equivalent to the phase of the Moon (at high latitudes the changing day length gives some indication, but this is much more

[3] However, Steele (*Journal of Astronomical History and Heritage*, **16** (2013) 250-260) makes a sensible case for the independence of Chinese and Babylonian astronomical systems, most notably that Babylonian positional astronomy is defined relative to the ecliptic, and Chinese relative to the equator.

difficult to quantify than the lunar phase, and varies more gradually). Therefore, pure solar calendars are less common in antiquity than lunisolar calendars.

There is some evidence that the pre-classical Greeks used a solar calendar. The famed ancient poet Hesiod (regarded by the Greeks as second only to Homer) wrote a poem called *Works and Days* which is essentially a farmer's almanac. Hesiod uses heliacal risings and settings to mark stations in the farming year: the Pleiades and Hyades clusters, the bright stars Arcturus and Sirius, and the easily recognised constellation Orion are all cited. This may, of course, have been an unofficial "folk" calendar (a literary equivalent of things like "Ne'er cast a clout/till May be out") rather than an official issue; Hesiod is very early, around 700 BC, and the Greek city-state system was not yet in existence. The recorded Greek calendars are lunisolar, with intercalation rather haphazardly determined by local officials (it is not clear whether Meton's suggestion of using his 19-year cycle was really adopted).

Our own calendar is a pure solar calendar which evolved from a lunisolar calendar as a result of official frustration. Intercalations in the Roman lunisolar calendar were the responsibility of the Pontifex Maximus ("High Priest"), who by the later Republic was not a career ecclesiastic, still less an astronomer, but in fact a politician. Roman politicians were much too busy furthering their own careers to worry unduly about minor matters such as maintaining the calendar, and by the pontificate of Julius Caesar it was in a complete mess. Caesar (somewhat surprisingly) decided to sort it out, consulted experts from Alexandria, and abandoned the old lunisolar calendar in favour of a pure solar calendar, retaining the division into 12 months but removing any attempt to stay synchronised with the Moon. The resulting **Julian calendar** is essentially our modern system, with months of 30 or 31 days (except February; note the difference from the 29- or 30-day true lunar months) and a leap year every fourth year. This gives a year length of 365.25 days, a little too long; by the 15th century this had become noticeable, and the Catholic establishment had decided that something must be done (the discrepancy was affecting the calculation of the date of Easter, and was therefore relevant to Church business). An initial attempt by Pope Sixtus IV was forestalled by the early death of his chosen expert, Regiomontanus (a very good choice, see later); the reform was finally approved by Pope Gregory XIII in 1582 (though the fact that the Reformation had taken place in the interim meant that it was not adopted in Britain until 1752). The necessary modification was extremely simple: century years are not leap years unless they are divisible by 400 (so 2000 was, but 2100 will not be). This removes 3 leap days from every 400 years, reducing the effective length of the year from 365.25 days to 365.2425—an excellent approximation to the true value of 365.2422 (in principle, we need to remove an extra day every 3300 years or so).

The Persian solar calendar, dating from the 11th century AD, is a better example of a solar calendar than the Gregorian calendar: it begins on the vernal equinox, and has six 31-day months followed by five 30-day months, and a twelfth month which has either 29 or 30 days depending on the time of the vernal equinox. Though the choice of 12 months suggests a lunisolar ancestry, the "long months followed by short months" structure is a better match to the motion of the Sun than the alternating pattern of our months (which is presumably a fossil from the lunisolar era); the Sun appears to move more slowly from March to September than from September to the following March, owing to the Earth's eccentric orbit. Because the Persian New Year is astronomically defined, there is no formal leap-year system (though of course a 366-day year will happen on average every four years).

3.8 Time

For early societies, exact measurement of time was probably rarely necessary in everyday life. A rough estimate of the time of day can be obtained from the height and direction of the Sun; water-clocks, graduated candles, hourglasses and such can be used to measure duration. Astronomers and astrologers would, however, require more precise measurements of time of night.

The natural way to measure time at night is by the rising of specific stars. The Babylonians divided the stars of the ecliptic (originally the stars in the path of the Moon, which isn't *quite* the ecliptic) into twelve constellations, and later refined this into twelve 30° segments of the ecliptic (the **signs of the zodiac**). The choice of 12 may reflect the 12 lunar months in a common lunar year. Defining time by the rising of successive zodiacal constellations then gives you 12 "double-hours" in a 24 hour day (in fact this would be a sidereal day, not a solar day, but the difference is only 4 minutes, so within a single night it would not be significant). The Babylonian *beru* was indeed a 2-hour period, divided into 30 *ush*. The fact that the Babylonians used the same word, "ush", for 1/360 of a circle and 1/360 of a day suggests that this reasoning is on the right lines. The Chinese had a similar system of double-hours.

We have retained the 60-fold divisions of time that derive from the Babylonians, but have divided our day into 2×12 hours (not, originally, 24, but specifically 12 daylight hours and 12 night hours—except at the equinoxes, day and night hours would be of different lengths) rather than 12 *beru*.

The division of the night into 12 hours appears to originate in Egypt. The Egyptians divided the celestial equator into 36 10° segments usually called (by us, not them) **decans**. The interval between risings of successive decans is about 40 minutes, so a 12-hour night would involve the rising of 18 decans. However, the Egyptians definitely seem to have taken the view that a night comprises 12, not 18, decan-rises [4]: tables listing these have been found painted on the interiors of coffins (don't ask me why the dead needed to know the time, but presumably there was some good reason in Egyptian mythology).

Extending the concept of precise time-keeping from the night to the day requires the use of the Sun. The Sun's height and direction both vary through the day, but the direction is the easier variable to use. The shadow of a vertical stick will provide a usable clock and a compass: the shadow is at its shortest at local noon, when it is also pointing due south. Once south has been fixed using the shadow length, a scale can be laid out to read off the time from the shadow direction.

This system can be made more elegant by inclining the stick to point at the celestial pole: this cancels out the variation in shadow length and allows the scale to be circular and the hours to be equally spaced. This is a standard sundial. Sundials measure what we now call *local apparent solar time*, which differs from *local mean solar time* by up to 15 minutes. As we have seen, both the Babylonians and the Hellenistic Greeks were well aware of the variations in the Sun's apparent motion, as were the Chinese, and Ptolemy corrected times from apparent to mean solar time in his writings.

[4] Quite why 12 decan-rises, which would take only eight hours, were regarded as defining the entire night is not clear; Wikipedia claims that the three either side were defined as twilight, but cites no source for this reasonable interpretation.

3.9 Alignment, Navigation and Mapping

Many ancient societies have placed great emphasis on accurate alignment of buildings, usually for religious purposes. We have discussed the astronomical/seasonal alignments of megalithic monuments earlier, and the accurate NSEW alignment of the Giza pyramids is also well known.

Islamic astronomers were very concerned with the process of calculating the **qibla**, or orientation to Mecca, enjoined on the faithful for prayer. (This is a much more difficult problem than direct astronomical alignments or NSEW, and contributed greatly to the Islamic astronomers' development of the techniques of spherical trigonometry: an example of an advance in astronomical techniques which is directly motivated by religion.)

Nowadays most people are aware that the Pole Star, α Canis Majoris, can be used at night to find North. However, because of precession, this was not the case 2000 years ago: the Hellenistic astronomers had no direct polar marker at all. Earlier, around 3000 BC, the star Thuban was a good polar marker, though considerably fainter than Polaris.

In the absence of a good polar marker, there are various ways of using the stars to establish the cardinal directions. Thurston suggests constructing an artificial level horizon (e.g. a semicircular wall; one could use a water channel in the top of the wall to ensure that it is accurately level). If you mark the points on this wall at which a star rises and sets, the line joining the rising and the setting is oriented east-west, and the line from the observer's station to the mid-point between rising and setting (defined either on the straight line or along the wall) is aligned north-south. Once the east-west direction is defined, further observations can identify a star which is on the celestial equator, and therefore rises due east and sets due west; this star can then be used as a standard alignment marker. If this standard marker is used over generations without recalibration, precession will cause the alignments to change over time relative to true east-west, and this could in principle be used to identify the marker. On the basis of the pattern of offsets from true east-west, Haack (*Archaeoastronomy* **7** (1984) S119-S125; reference in Thurston) suggests that the Giza pyramids were aligned using β Scorpii.

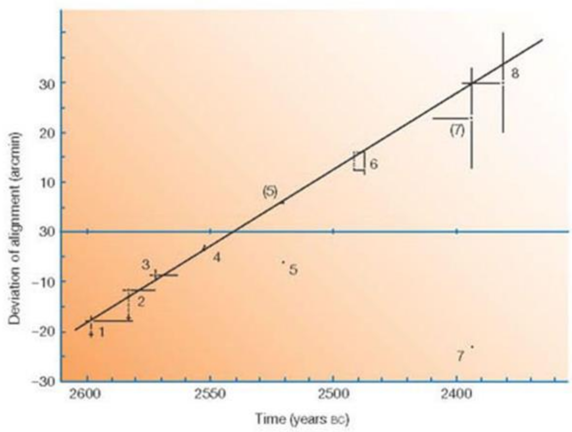

Figure 3.2: Alignment of pyramids from 2600 to 2300 BC, from K Spence, Nature ***412*** *(2001) 699-700. The numbers correspond to Meidum (1), the Bent Pyramid (2), the Red Pyramid (3), Khufu (4) [the Great Pyramid], Khafre (5), Menkaure (6), Sahure (7) and Neferirkare (8). The bracketed 5 and 7 correspond to flipping the sign of the deviation: the unbracketed negative values are the true measurements. See text for the explanation of how such a flip can arise naturally from the proposed method.*

The problem with this method is that observations of rising and setting stars are easily thrown off by a non-level horizon. An alternative method, suggested by Kate Spence (*Nature* **408** (2000) 320-324), relies on observing stars at transit (or culmination), at which time they are due north or due south depending on their declination. Specifically, she suggests that the Egyptians made use of two circumpolar stars, β Ursae Minoris and ζ Ursae Majoris, which at that date lie 12^h apart in right ascension: the upper culmination of one then coincides with the lower culmination of the other, so that the two stars should be aligned vertically above each other in the sky (verifiable observationally using a plumb line). This direction is then true north. The

Egyptians are known to have been interested in circumpolar stars (which they regarded as "immortal", since they equated setting with death), so it is very plausible that they would wish to use these stars for ritual alignments. Many archaeologists also believe that four narrow shafts in Khufu's pyramid are aligned on the culminations of important stars: the candidate star for one of these alignments is again β Ursae Minoris. As with the β Scorpii hypothesis, the supporting evidence adduced by Spence is the good fit of the offsets from true north to the error introduced by precession; Spence demonstrates, as shown in figure 3.2, that a good fit can be obtained to the offsets of eight pyramids if a slight adjustment is made to the standard Egyptian chronology. An interesting effect, which would not occur with the preceding method, is that two of the pyramids appear at the mirror image of their predicted error: this would happen if the surveyors for those pyramids swapped the two stars over, using the upper culmination of ζ UMa instead of its lower culmination (β UMi is slightly brighter than ζ UMa, which might explain why most pyramids were aligned with it as the upper star). Spence argues (*Nature* **412** (2001) 700) that her proposed method is not only easier to use, but also supported by pictorial and textual evidence from later tombs.

The Egyptians do not seem to have developed the concept of terrestrial longitude and latitude: with an essentially linear country, they probably saw no practical need for such a thing. Both the Greeks and the Chinese did grasp the idea of latitude, and its connection to the height of the Sun at noon, but they interpreted these observations in rather different ways.

The Greeks developed the concept of a spherical Earth quite early: their arguments included the curved shape of the Earth's shadow on the Moon during a lunar eclipse, and the "hull-down on the horizon" observation whereby the mast and rigging of a distant ship come into view while the hull is still below the horizon. They then interpreted the change in height of the Sun as a consequence of the shape of the Earth. In the Hellenistic era, **Eratosthenes** used this variation to estimate the circumference of the Earth, by observing that on a certain date the Sun was directly overhead at a place called Syene (this was demonstrated by noting that its reflection could be seen at the bottom of a deep well), but cast a small measurable shadow in Alexandria, which was 5000 stadia almost directly due north. He therefore deduced a value of 252000 stadia for the Earth's circumference.

Opinion is divided as to the accuracy of this result—it is not clear which of several definitions of the stadion he was using—but it is certainly of the correct order of magnitude. Thurston observes that Eratosthenes may have been using (unknowingly) circular reasoning: it is very possible that the distance between Syene and Alexandria was initially calculated not by surveying but by the same sort of geometry applied in reverse! However, this merely credits someone else with the original calculation: it does not invalidate the argument that the Hellenistic Greeks had a good idea of the approximate shape and size of the Earth, and grasped the concept of using the Sun's altitude at noon to measure latitude.

The Chinese, in contrast, sometimes interpreted the changing altitude of the Sun as a parallax effect on a flat Earth, and therefore used it to estimate the height of the Sun above the ground. This is an example of the way in which experimental data can be interpreted differently depending on your underlying theory. In the Greek picture, the Sun is at a very large distance, the Earth is spherical, and the distance between the two observing sites (measured along the curved surface) is $R\Delta\theta$; if you know the distance, you can read off the Earth's radius R. In the Chinese picture, the difference in shadow lengths is caused by the finite height of the Sun above the flat Earth, and if you know the distance you can read off the Sun's height h. Given that the Chinese pictures of the relation of Earth and sky did in fact involve a spherical or hemispherical

Earth, it is slightly surprising that they were happy with this calculation, but it appears that Chinese abstract geometry was not as advanced as that of the Greeks.

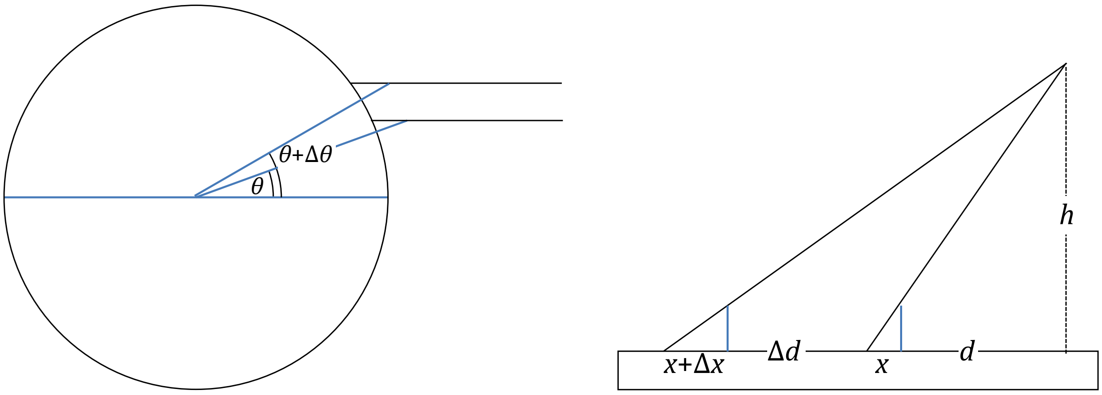

Figure 3.3: shadow lengths. Left, the Greek picture: the difference in shadow lengths determines $\Delta\theta$, and knowing the distance Δd between the gnomons allows you to calculate the radius R of the Earth: $\Delta d = R\,\Delta\theta$. Right, the Chinese picture: if the height of the gnomon is l, then by similar triangles $x/l = d/h$, and $(x+\Delta x)/l = (d+\Delta d)/h$. Hence $\Delta x/l = \Delta d/h$, and knowing Δd allows you to calculate the height of the Sun, h.

The astronomical determination of latitude is straightforward with the right mental picture. Determining longitude is a much more difficult task. The simplest solution, if you accept a spherical Earth, is to compare the local solar time with a reference solar time; this is what was done in modern times until the advent of GPS. However, in the absence of high-speed communications, this requires access to a reliable chronometer (so that the remote site *knows* the reference solar time). In antiquity, there were no sufficiently reliable clocks (transporting a water-clock on an overland surveying journey, let alone a sea voyage, is not practical). It is in principle possible to determine the reference time using predictable astronomical events or quantities; you then use the Sun to determine local solar time, and read off the longitude. Hipparchos suggested using the timing of a lunar eclipse as the astronomical fixed reference: this does work, but is only practical for fixed locations on land, because of the rarity of lunar eclipses. It is therefore useful for surveying and map-making but not for navigation.

In his *Ephemerides* of 1474, Regiomontanus suggested using a table of the angular distance at a given time between the Moon and a series of suitable bright stars as a reference "clock". This **method of lunar distances** became the standard technique for land surveying. It was developed for use at sea by the Astronomer Royal **Nevil Maskelyne**, who published the first Nautical Almanac in 1766, and briefly used during the late 18^{th} century, before being superseded by reliable marine chronometers.

To summarise: alignment, map-making and navigation all relied until modern times on astronomical observations. Alignment is fairly straightforward, and most ancient civilisations have produced accurately aligned monuments (what they are aligned *with* of course depends on the particular culture). Latitude determination is straightforward provided you have the correct mental model (attempts to determine "latitude" with an assumed flat Earth and Sun at finite distance will eventually go wrong). Longitude, with its need for a reliable absolute standard of time, is an extremely difficult problem; marine longitude determination (from a moving platform, at an arbitrary time—you can't schedule your trading voyages according to the lunar eclipse cycle!) is the worst of all, and was only solved satisfactorily in the mid-18^{th} century (the method of lunar distances could in principle have been used from the 15^{th} century onwards, but the instrumentation available made the necessary measurements difficult—the cross-staff is not the instrument of choice for use on a tossing ship in mid-Atlantic!—and the calculations required are not trivial.

3.10 Summary: Earth and Sky

Before the widespread use of street lighting, the various periodic changes in the night sky—the rising and setting of Sun, Moon, planets and stars, the changing phase of the Moon, and the changing visibility of different constellations over the course of the year—would have been very noticeable to anyone who looked at the night sky. It is not surprising that ancient civilisations in need of a calendar system generally made use of these periodic changes. This would have led naturally to the desire to keep records of the relevant astronomical markers in order to improve the calendar system, leading eventually to the development of quite sophisticated systems of positional astronomy.

The various Eurasian systems of astronomy and calendrics—Greek, Babylonian, Indian and Chinese—have a number of similarities, such as the use of the Metonic cycle for calendar regulation (and indeed the use of lunisolar calendars in general). Some of these traditions are surely not completely independent: there were Greek city-states on what is now the Turkish coast from early times, Indian astronomy after the 4th century BC is clearly influenced by Hellenistic models, and there are records of Indian astronomers resident in China. On the other hand, there are enough differences (for example, the Chinese system of constellations, which is not the same as the Greek/Babylonian set) to imply that there were long periods of independent development in between the points of contact. As all civilisations are looking at the same sky, it is entirely possible that convenient common multiples such as the Metonic cycle and the saros could be discovered independently.

The Mesoamerican calendar system, used by most of the civilisations of central America although nowadays mostly associated with the Maya, is sufficiently different to lend indirect support to the idea of mutual influence among the Eurasian cultures (just because it proves that ancient societies do not automatically hit on the idea of a lunisolar calendar). This system has a "civil year" or *haab* of 365 days (which will drift slowly out of synchronisation with the solar year, because of the lack of a leap-year mechanism) and a religious cycle, the *tzolkin*, of 260 days. Combining the *haab* and the *tzolkin* gives a "Calendar Round" of about 52 years—the lowest common multiple of 260 and 365 is 18980 days = 52×365 = 73×260. To deal with periods longer than the Calendar Round, Mesoamerican cultures used the *Long Count*, which is essentially just a count of days in base-20 place-value notation (except that the "hundreds" place is 360 days, 18×20, instead of 400—presumably because 360 corresponds roughly to the solar year—so that the next two places are 7200, 20×18×20, and 144000, 20^2×18×20, days) from a start date of August 11, 3114 BC (which is millennia before the first attested Long Count date but is apparently a convenient starting date for a number of astronomical cycles; note that the astronomical Julian date also has its zero placed way before any usage of the system, in either 4713 or 4714 BC depending on whether you extrapolate back using the Julian or Gregorian calendar). Although the Maya in particular were accomplished astronomers who kept careful records of astronomical phenomena (especially the planet Venus), none of these systems has any strong links to astronomy (it has been suggested that the 260-day *tzolkin* might have originated as the time between successive zenith passages of the Sun at latitude 15° N—but it's also been suggested that it derives from the length of the human pregnancy, so this cannot be taken as proof of an astronomical connection!).

Even if the original motivation concerns calendar calculations, careful observations of the motion of the Sun, Moon and stars will inevitably lead to the recognition of significant directions: the stars appear to rotate about a stationary point in the sky (which, owing to precession, may or may not be marked by a star, see figure 3.4); the Sun reaches its highest altitude when it is due South (in the northern hemisphere), and the points on the horizon where it rises and sets move north and south according to the season. These directions may develop religious significance, as appears to have happened in the culture that built the megalithic monuments in Britain, and astronomical observations can be used to align structures in religiously meaningful ways.

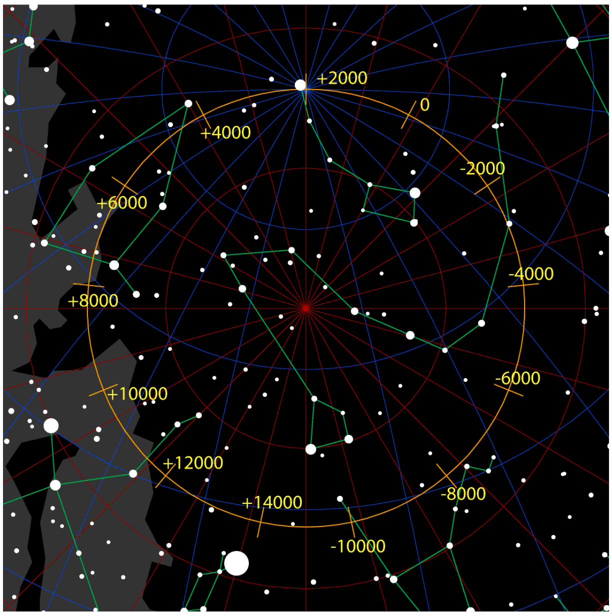

Figure 3.4: precession. The circle shows the path of the North Celestial Pole from 10000 BC to AD 14000. Note that most of the time there is no "pole star" as such: Polaris is a good pole star now, Thuban was a good if rather faint one around 3000 BC, and γ Cephei will be fairly close around AD 4000.

Finally, any culture that occupies a territory with significant north-south extent will notice that the maximum altitude attained by the Sun and the stars varies according to the observer's latitude. This provides such cultures with a tool for determining latitude, which can be used to assist in mapping and navigation, and—given the correct mental model, and some grasp of geometry—can be applied to determine the radius of the Earth. This last calculation implies a progression from simply developing mathematical tools to predict celestial phenomena and regulate calendars to constructing mathematical models of the planetary system, a step which in the European tradition is associated with the ancient Greeks, and more precisely with the eastern Mediterranean Greeks of the Hellenistic period, who had the opportunity to combine Babylonian record-keeping and mathematical notation with Greek formal geometry. This is the subject of the next topic.