PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Introduction to Some Basic Astronomical Concepts

R. R. Newton

Phil. Trans. R. Soc. Lond. A 1974 **276**, 5-20
doi: 10.1098/rsta.1974.0006

Phil. Trans. R. Soc. Lond. A. **276**, 5–20 (1974) [5]
Printed in Great Britain

I. ASTRONOMY IN ANCIENT LITERATE SOCIETIES

Introduction to some basic astronomical concepts

By R. R. Newton
Applied Physics Laboratory, Johns Hopkins University
Silver Spring, Maryland 20910, *U.S.A.*

In modern celestial mechanics, the goal is to construct a coordinate system in the heavens, and then to develop theories that will allow us to calculate the coordinates of the Sun, Moon and planets as functions of time. The goal of astronomy among ancient peoples was probably quite different; it was probably to develop methods of dealing with certain appearances in a coordinate system based upon the observer's horizon. These appearances include eclipses, the positions at which celestial objects rise and set, and methods of telling both the time of day and the time of the year by observations of risings and settings. This paper is concerned with discussions of these appearances. It deals with theories of motion only to the extent needed in discussing the appearances.

1. Introduction

In studying the place of astronomy in the ancient world, it is necessary to have some knowledge of the basic astronomical phenomena that ancient peoples observed. The purpose of this paper is to outline some basic elements for the sake of those who have not had a previous opportunity to study them. Much if not all of the paper will be familiar to many readers.

The first task is to establish coordinate systems which can be used in describing the important phenomena that concern us. Some coordinate systems are fixed with respect to an observer on the Earth's surface. Others are fixed in the heavens; that is, the stars appear to be at constant positions in these systems except for some very slow motions. After we establish the necessary coordinate systems and the relations between them, we turn our attention to three important classes of problem.

The first class concerns the risings and settings of celestial bodies. We shall be concerned with the places on the horizon at which rising and setting occur. We shall also be concerned with the times of rising and setting, both with regard to time of day and time of year.

The second class concerns the motions of the Sun, Moon, and planets with respect to the stars. Only the most basic elements of motion will be discussed here, in modern terms. The terms in which planetary motion was discussed by ancient peoples will be left as a matter for the specialists in those subjects.

Finally, we shall be concerned with the theory of eclipses, since eclipses often played a dramatic part in the thinking of ancient peoples.

In all the discussions, I shall treat the Earth as if it were an exact sphere. I shall also neglect various other small effects such as refraction.

2. Earth-fixed and star-fixed coordinates

The coordinate system that was most obvious to an ancient observer was probably the one based upon his own horizon, which we can assume to be level for present purposes. Suppose that an observer is looking north from the point marked O in figure 1, so that he sees half of his horizon along the arc WHNE. Z is his zenith and N is the direction of north. Thus the plane ZON is part of his meridian plane. Other vertical planes are drawn in lightly.

Suppose that the observer sees an object at the point Q. He locates it by first drawing the vertical plane HOZQ through the point and by measuring the angle HON in his horizon plane. HON is called the azimuth of the point Q. HON is usually considered positive when it measured clockwise from N, thus the azimuth of Q is a negative number as figure 1 is drawn. Various conventions about the starting-point and the direction of measuring azimuth have been used in the past. Azimuth will be denoted by the symbol A.

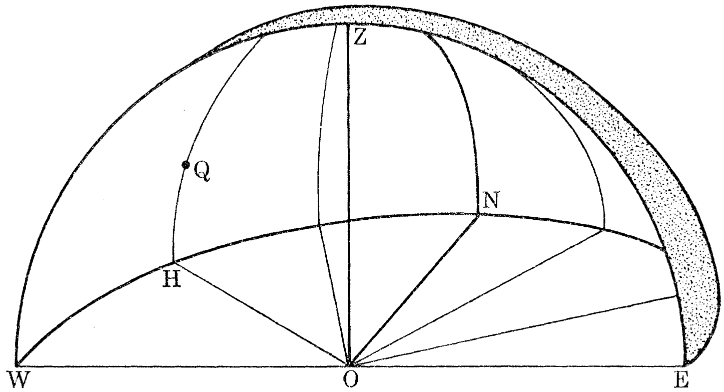

FIGURE 1. The coordinate system of azimuth and elevation.

The observer completes the location of Q by measuring the length of the arc QH, which is called the elevation a of the point Q.

I shall use the term anthropocentric for the coordinate system of figure 1, in order to emphasize that it is based upon the perception of a specific individual at a specific spot. The term topocentric is probably more common, but it removes attention from the individual.

We need the distance from the observer to the object in order to complete the anthropocentric system. For astronomical bodies, the distance is usually so large that its value is unimportant, and the distance coordinate will receive little attention in this paper.

Let us now turn to star-fixed coordinates. I shall assume that the reader knows that the Earth spins on an axis. The points at which the axis pierces the surface of the earth are called the north and south poles. The points at which the axis pierces the celestial sphere, that is, the spherical surface upon which the stars and planets seem to lie, are called the north and south celestial poles. With the aid of these poles, we can set up a coordinate system in the heavens that is analogous to the longitude–latitude system used on earth.

This coordinate system is illustrated in figure 2. In this figure, we are looking at a hemisphere as seen from inside. The figure shows circles that look much like the lines of longitude and latitude shown on terrestrial globes. The quantity analogous to longitude is called the right ascension α, and the quantity analogous to latitude is called the declination δ. Declination, like latitude, is positive north of the equator and negative south of it. Right ascension, like longitude, increases as we go eastwards. There is one main difference between figure 2 and the corresponding figure of a terrestrial hemisphere. In representations of the Earth, we show how it looks from the outside, but in figure 2 we are looking from the inside. Thus, with north at the top, east is to the left rather than to the right in figure 2.

Except for small effects that can be neglected here, a star has a fixed position in the coordinate

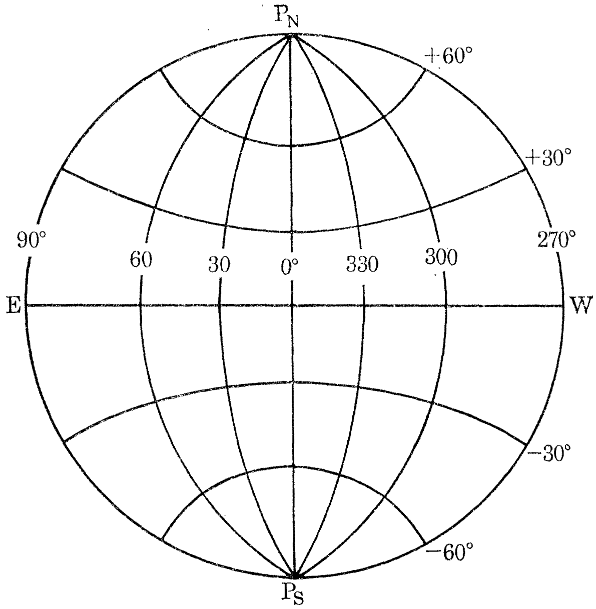

FIGURE 2. The coordinate system of right ascension and declination.

system of figure 2. Obviously a star does not have a fixed position in the system of figure 1. Our task is now to relate positions in the two systems.

In establishing the connexion between the coordinate systems, I am going to assume that the observer's horizon plane passes through the centre of the Earth instead of being tangent to the Earth. The effect of this assumption is shown in figure 3. In this figure, we see the observer's actual horizon drawn tangent to the Earth's surface. The parallel plane through the centre of the Earth can be called the translated horizon. It is clear that the translation of the horizon plane does not change the azimuth of any point. The figure shows how the elevation angle changes from one horizon to the other. The difference between the two elevation angles is called the parallax.

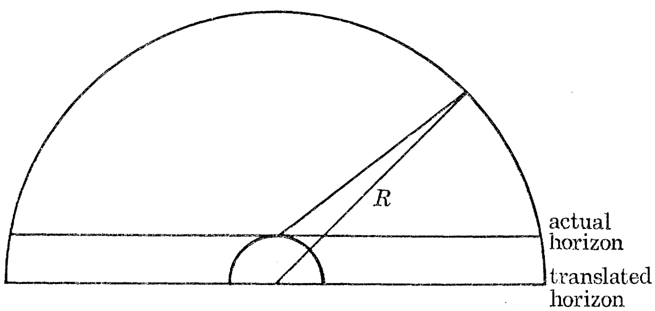

FIGURE 3. The phenomenon of parallax, shown as a function of distance R and of elevation angle.

The parallax depends upon the elevation of the point considered, and it is obviously greatest for a point on the horizon. The parallax for a point on the horizon is called the horizontal parallax, which clearly depends upon the distance R of the point from the centre of the Earth. For the Moon, the horizontal parallax is about 1°. For every other body that will concern us here, the horizontal parallax is much less than 1′. I shall neglect parallax in the remaining discussion of coordinate systems. This is equivalent to taking the translated horizon as the actual horizon.

An observer can establish an Earth-fixed coordinate system for describing the heavens that, like the celestial system of figure 2, is based upon the poles. The appearance of this system will necessarily depend upon the observer's latitude. Its appearance for an observer on the equator is shown in figure 4.

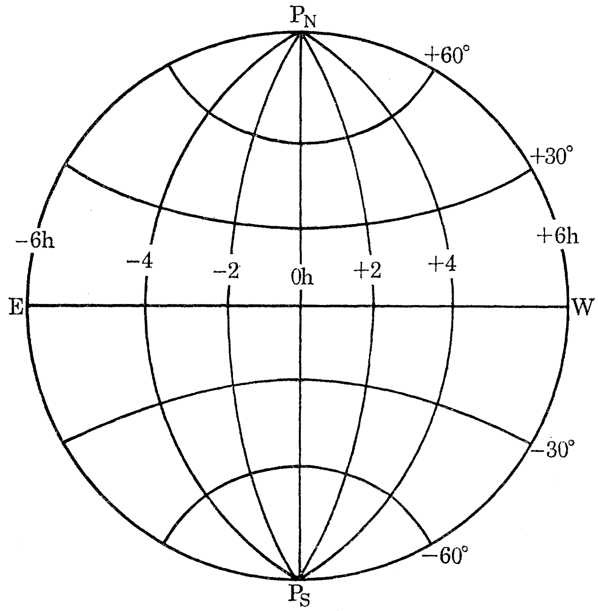

FIGURE 4. The coordinate system of hour angle and declination, for an observer on the Equator.

Suppose that an observer is lying on his back at the equator, with his head pointing toward the north. The celestial poles P_N and P_S are on the horizon. Planes parallel to the equator cut the celestial sphere in circles that are identical to the declination circles in figure 2, and we can continue to use the term declination for the corresponding coordinate. Planes passing through the poles cut the celestial sphere in circles that look like the circles of right ascension in figure 2. However, there are two main differences. The circles in figure 2 are fixed in the stars and rotate with respect to the Earth, while the circles in figure 4 are fixed to the Earth and rotate with respect to the stars. Secondly, we number the circles in figure 4 with the zero circle passing through the zenith and with positive values to the west.

The coordinate defined by the planes passing through the poles in figure 4 is called the hour angle h. It is shown in figure 4 measured in units of hours, while the right ascension in figure 2 was shown in units of degrees. The relation between the units is

$$1 \text{ hour of angle} = 15°. \tag{2.1}$$

Both right ascension and hour angle are measured in hours in most literature, but the use of degrees is probably increasing.

Now suppose that the observer is in the northern hemisphere at latitude ϕ, and let us suppose that he is standing looking toward the north as he was in figure 1. Then the elevation angle of the north celestial pole, namely the arc NP in figure 5, equals ϕ. The declination circles for $\delta \geqslant 90° - \phi$ are above the horizon for their full circuit. If the observer turns toward the south, he finds that the south celestial pole is below the horizon by the angle ϕ; the declination circles that lie within the angle ϕ from the south pole never rise above the horizon. Intermediate declination circles are visible in part but not in full.

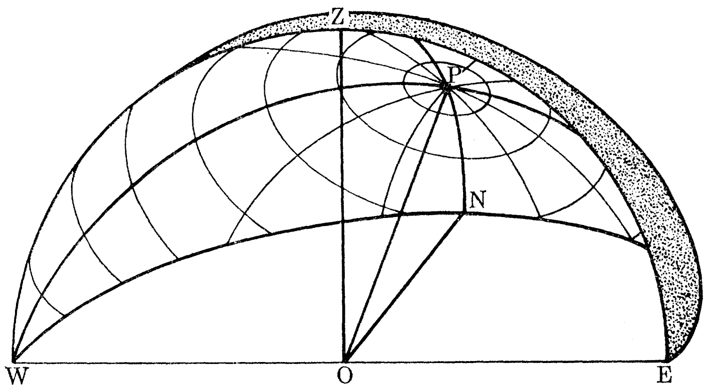

FIGURE 5. The coordinate system of hour angle and declination, for an observer in the northern hemisphere.

The hour-angle arcs radiate from the pole in figure 5. The hour angle is zero along the arc PZ. It is -6h along the arc PE and $+6$h along the arc PW. Along the arc PN, it can be called either -12h or $+12$h, whichever is convenient. Some arcs for other values of the hour angle are drawn in lightly in figure 5.

A star or other celestial object with a declination δ seems to travel around the Earth on the corresponding declination circle, like a bead sliding on a string. Thus the stars near the north pole never set for the observer in figure 5, although they may become invisible in the daytime because of the brightness of the sky. Stars near the south pole never rise. All stars in between rise, remain above the horizon for some time, and then set. Since the declination circles in figure 5 do not intersect the horizon at right angles, an object does not rise straight from the horizon. Instead, it travels southward as it rises and, similarly, it travels northward as it sets. Only at the equator do objects rise vertically.

I shall take up the formal relations between the various coordinate systems in the next section. After the formal relations have been established, I shall take up rising and setting problems.

3. TRANSFORMATIONS OF COORDINATES; TIME

The azimuth-elevation system in figure 1 and the declination-hour angle system in figures 4 and 5 are fixed with respect to the Earth. Hence the transformation between the systems does not involve the time, but it does involve the latitude ϕ of the observer. The transformation from (a, A) to (δ, h) is

$$\left.\begin{aligned} \delta &= \arcsin\{\sin a \sin\phi + \cos a \cos A \cos\phi\}, \\ h &= \arctan\{-\cos a \sin A/(\sin a \cos\phi - \cos a \cos A \sin\phi)\}. \end{aligned}\right\} \tag{3.1}$$

The inverse transformation from (δ, h) to (a, A) is

$$\left.\begin{aligned} a &= \arcsin\{\sin\delta \sin\phi + \cos\delta \cos h \cos\phi\}, \\ A &= \arctan\{-\cos\delta \sin h/(\sin\delta \cos\phi - \cos\delta \cos h \sin\phi)\}. \end{aligned}\right\} \tag{3.2}$$

Since δ and a are restricted to lie between -90 and $+90°$, there is no ambiguity in the inverse sine function.

The transformation to the right ascension–declination system does involve the time. In order to derive this transformation, we note that right ascension α increases to the east while hour angle h increases to the west. Therefore, at any instant of time, the sum $\alpha+h$ is the same for all bodies. At a later instant, α is still the same as it was earlier, for a fixed star, but all stars have moved to the west as seen from Earth. Thus h increases with time for a given star. In other words, the sum $\alpha+h$ is a monotone in creasing function of time, but the sum has the same value for all stars at any specific instant.

The function of time defined by the sum $\alpha+h$ is called local sidereal time, which I shall abbreviate as L.S.T. Thus:

$$\text{L.S.T.} = \alpha+h. \tag{3.3}$$

Although the value of L.S.T. does not depend upon the star used, it does depend upon the terrestrial longitude of the observer. The value of L.S.T. for 0° longitude is called Greenwich sidereal time, often abbreviated G.M.S.T. The 'M' in G.M.S.T. stands for 'mean'. In very accurate work, it is necessary to distinguish between 'mean' and 'apparent' sidereal time, but this distinction need not concern us in the study of ancient astronomy.

The adjective 'sidereal' refers to the stars and in conversation we often say that sidereal time is the time measured by the fixed stars. This is not quite correct. From equation (3.3) we see that local sidereal time is equal to the hour angle of the point whose right ascension is zero. This point is called the vernal equinox, or the first line of Aries. The vernal equinox moves slowly westward with respect to the stars and makes one complete circuit in about 26000 years. Thus, in about 26000 years, the time that would be measured by the stars will be 1 day greater than sidereal time as defined by equation (3.3). This fact also means that it is not quite correct to call the coordinate system (α, δ) of figure 2 a star-fixed system. However, I shall ignore this and continue to call figure 2 a star-fixed system.

Since δ is the same in both the star-fixed and Earth-fixed systems of figures 2, 4 and 5, the transformation between the systems involves only α and h. Equation (3.3) is thus the desired transformation.

Equation (3.3) also furnishes a practical means of measuring time. We have only to measure the hour angle of an object whose right ascension α is known. It is possible to mount a telescope or other sighting tube on gimbals so that one axis is parallel to the Earth's axis, making one plane of rotation parallel to the equator. With such an 'equatorial' mount, the hour angle can be measured directly; however, ancient and medieval astronomers do not seem to have used this method much.

The method of measuring time astronomically that was most used, at least if we judge by the surviving data, was that of measuring the elevation of a star. Since δ is known for the star, we find h from a by eliminating A from equations (3.1). This is in fact just what we did in deriving the first of equations (3.2), so

$$\cos h = (\sin a - \sin\delta \sin\phi)/\cos\delta \cos\phi. \tag{3.4}$$

This is not useful when h is near 0, because $\cos h$ changes too slowly there. However, the astronomers seem to have applied this method only to stars fairly near the horizon. We should note that $\cos h$ has the same value for positive or negative h, so that it is necessary to know whether the star used was in the east or the west.

It would work just as well to measure the azimuth, although this does not seem to have been done as often as measuring the elevation. In order to find h from A, we solve the second of equations (3.2) for h. This requires converting the equation into a quadratic equation for $\sin h$ (or $\cos h$), solving the quadratic, and then finding the inverse function. The cumbersomeness of this method may have inhibited its use.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The time commonly used in daily life is not sidereal time but solar time, based upon the apparent daily rotation of the Sun. Until 1925, astronomers for many centuries and perhaps millennia had defined solar time to be the hour angle of the Sun. Since the hour angle of the Sun is zero at midday, this meant that astronomers began their day at midday, which put them at variance with most of the population. Various peoples have had various conventions for beginning the day, such as sunrise, sunset, or midnight, but few if any (except astronomers) have ever begun their day at the time that we call noon. Accordingly, the definition of solar time used by astronomers was changed in 1925, and it is now defined as 12h plus the hour angle of the sun.

Solar time as defined depends upon the longitude of the observer, and its value for the meridian of Greenwich is called G.M.T. (Greenwich mean time). As we did with sidereal time, we shall ignore the significance of the adjective 'mean' inserted into this definition. G.M.T. is almost synonymous with the quantity called 'universal time'.

G.M.T. and G.M.S.T. can be related as soon as we know the right ascension of the Sun at any time during the year. We can also calculate one from the other simply but tediously by noting that the times are equal by the old definition when the Sun is at the vernal equinox and that they are equal by the new definition when the Sun is at the autumnal equinox. During a tropical year, the stars make one more apparent revolution than the Sun; that is, if there are D solar days in a tropical year, there are $D+1$ sidereal days. Since D is approximately 365.2422,

$$\begin{aligned}\text{sidereal time/solar time} &= 366.2422/365.2422 \\ &= 1.002738,\end{aligned} \tag{3.5}$$

if we reckon both times from a common epoch. Ephemeris publications contain tables that aid in applying equation (3.5) at any time during the year.

In sum, one way that time appears in astronomy is through the transformation between Earth-fixed and star-fixed coordinates. If an astronomer wishes to measure time during daylight, he does so by finding the hour angle of the Sun in his anthropocentric coordinate system. If he wishes to measure time during the dark hours, he can do so by finding the hour angle of a star and thence finding sidereal time from equation (3.3). He usually converts this time to solar time by using equation (3.5) or the equivalent.

One of the oldest ways of using this idea was to use the special case in which $a = 0$, that is, to use celestial objects on the horizon. This brings us to the subject of risings and settings.

4. Risings and settings

From figure 5, we see that an object never sets if $\delta > 90° - \phi$, and that it never rises if $\delta < \phi - 90°$, for an observer in the northern hemisphere. Symmetrical relations apply in the southern hemisphere. Objects with intermediate declinations do rise and set, and we see from figure 5 that the azimuths of the rising and setting points depend only upon the declination. Thus a given star always rises and sets at the same points on the horizon.

This fact furnishes a simple way of measuring declination if we know our latitude, and conversely. We measure the azimuth A of an object either at rising or setting. Then, in order to relate declination and latitude, we set $a = 0$ in the first of equations (3.1), obtaining simply

$$\sin \delta = \cos A \cos \phi. \tag{4.1}$$

If the horizon is not level, we must use the correct value of a at the rising or setting point, which gives us a slightly more complex relation.

The hour angle at rising or setting is also a unique function of declination in a given anthropocentric system. The relation between declination and hour angle, when $a = 0$, is a special case of equation (3.4):

$$\cos h = -\tan \delta \tan \phi. \tag{4.2}$$

Risings and settings furnish a convenient way of measuring time. If we observe that a certain star, say, is about to set, we use its known declination to find h from equation (4.2). We then use its known right ascension to find sidereal time from equation (3.3), and we can then convert to solar time if we wish.

For a given star, the value of h at rise, say, is the same every sidereal day. That is, a star always rises at the same sidereal hour. The relation between the sidereal hour and the solar hour changes steadily throughout the year, however, so the solar time of rising also changes steadily with the time of year.

Conversely, if we note what star has just risen or is about to set at some solar time, sunset say, we can tell the season. This fact was important to farmers and others whose activities depend upon the season. A number of ancient works tell how to judge the correct times for planting by observing what stars are rising, say, at sunset.

We can use equation (4.2) to find how the length of the day varies with the seasons. In order to do so, we must use a fact that will be discussed in more detail below, namely that the declination of the Sun varies from about $-23°$ in midwinter to about $+23°$ in midsummer. Let δ_S denote the declination of the Sun at any time. We note that $\cos h$ has the same value whether h is positive or negative. Hence equation (4.2) furnishes two equal and opposite values of h; the negative value applies at rise and the positive value at set. Thus the length of the day is twice the value of h at sunset. Hence

$$\text{length of day} = 2 \arccos\,(-\tan \delta_S \tan \phi). \tag{4.3}$$

Let us evaluate this for a latitude of 50°. In midwinter, $\delta_S = -23°$, and $\cos h = +0.5$ rather closely. Hence h at set is 4h and the day is 8h long. At the equinoxes, $\delta_S = 0$, $\cos h = 0$, h at set is 6h, and the day is 12h long. In midsummer, $\delta_S = +23°$, $\cos h = -0.5$, h at set is 8h, and the day is 16h long. These figures are for the length of time the centre of the Sun is above the horizon, in the absence of refraction; other definitions of sunrise and sunset give slightly different values.

5. The ecliptic and the zodiac; unequal hours

Ancient observers noted that most of the objects in the heavens are fixed relative to each other, but that a few move in the heavens in a regular fashion. Thus they spoke of the fixed stars and the wandering stars or planets. In their terminology, there were seven planets: Sun, Moon, Mercury, Venus, Mars, Jupiter, and Saturn. Uranus, Neptune, and Pluto had not been discovered yet.

Within the accuracy that is possible for observations with the naked eye, the Sun moves in a plane that is fixed with respect to the stars. This plane is called the ecliptic. The ecliptic plane cuts the celestial equator in two points called the vernal equinox and the autumnal equinox; these points are also called the first line of Aries and the first line of Libra. The angle between the planes is called the obliquity ϵ. At noon on 1899 December 31, ϵ equalled 23° 27′ 08.26″ according to the best estimate. Because of perturbations of other planets upon the apparent motion of the Sun (that is, upon the orbital motion of the Earth in modern terms), ϵ varies with time. On the scale of geologic time, the variation is probably oscillatory. Within the relatively short scale of historic time, however, ϵ has been decreasing steadily at a rate of 46.845″ per century, to high accuracy.

At the present time, ϵ is closer to 23° than to any other integral value, and I used 23° earlier for illustrative purposes. In ancient historical times, however, ϵ was close to 24°. This fact is important in studies of ancient astronomy.

Since the ecliptic makes an angle with the equator, the Sun is north of the equator ($\delta_S > 0$) about half the time and it is south ($\delta_S < 0$) about half the time. The vernal equinox is the point at which δ_S changes from negative to positive and the autumnal equinox is the point at which δ_S changes in the opposite direction. The terms 'vernal equinox' and 'autumnal equinox' are used to mean either the points in the sky that were defined above or the instants in time when the Sun passes these points. The context usually makes it clear which meaning is intended.

After the Sun passes the vernal equinox, its declination continues to increase until it reaches the value ϵ. The instant and position of this event are called the summer solstice. The declination of the Sun then begins to decrease. When it is 0, we are at the autumnal equinox and when it is at its minimum $-\epsilon$, we are at the winter solstice.

I have already mentioned the fact that the equinoxes move with respect to the stars. This motion is caused by planetary and lunar perturbations, as is the change in the obliquity. Since the plane of the equator and the plane of the ecliptic both move, there is no simple way to describe the path of the equinoxes. However, the ecliptic moves much less than the equator and, for simplicity, we can say that the equinox moves westward along the (fixed) plane of the ecliptic; it makes a complete circle around the ecliptic in about 26000 years. This motion is called the precession of the equinoxes.

The planets other than the Sun (including the Moon) move in a narrow band in the heavens that is centred on the ecliptic and that lies within a few degrees of it. This band is called the zodiac. The zodiac is divided into 12 equal parts called signs. The names of the signs, in order going eastward, are: Aries, Taurus, Gemini, Cancer, Leo, Virgo, Libra, Scorpio, Sagittarius, Capricorn, Aquarius and Pisces. These names furnish part of the common terminology of astrology, which was not carefully distinguished from astronomy until modern times.

After the precession of the equinoxes was discovered in about the year 150 B.C., these names had to serve double duty. They were originally, and still are, the names of constellations, and the divisions of the zodiac were taken from these prominent constellations that lay within them. The vernal equinox, as we have said, is still often called the first line of Aries. Precession has carried the vernal equinox westward from the constellation Aries and it now lies within the constellation Pisces. However, by the time that the precession was discovered, the names were firmly established in the terminology used for specifying the positions of stars. Hellenistic astronomers such as Hipparchus and Ptolemy decided to retain the names to indicate equal sections of the zodiac, beginning with Aries at the vernal equinox and continuing eastward. That is, Aries, when considered as a sign of the zodiac rather than as a constellation, begins at the vernal equinox and includes 30° of the ecliptic circle going eastward. Taurus begins 30° east of the vernal equinox and ends 60° east, and so on.

The signs of the zodiac have thus become merely ways of helping to specify the astronomical

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

coordinate called celestial longitude, which will be defined in the next section. Before we turn to that definition, it is desirable to take up the matter of unequal or seasonal hours, which are closely related to the ecliptic.

The ecliptic cuts a great circle on the celestial sphere, and the horizon of any observer cuts another great circle which, for observers in temperate or tropical latitudes, never coincides with the ecliptic circle. Therefore half of the ecliptic is always above the horizon, if we restrict the discussion to observers outside the frigid zones.

Consider the Sun at the instant of sunset. It is in the ecliptic, at the point B, say. Let C denote the point on the ecliptic diametrically opposite to B. Since B is on the western horizon, by definition, C is on the eastern horizon. Since point B is the point of the next sunrise, except for the small angle that the sun moves along the ecliptic between sunset and the next sunrise, all the points on the ecliptic from C eastward 180° around to B rise during the night.

Suppose that we divide the ecliptic into 24 equal parts. Then 12 parts will rise during any night, regardless of the season and regardless of the observer's latitude. That is, if we keep time by observing the point of the zodiac that is on the horizon, instead of using the right ascension of a point on the horizon, we divide each night into 12 parts, called 'hours of the night', 'seasonal hours', or 'unequal hours'. We can also arrange for a sundial or other means of marking time to divide daylight into 12 parts, called 'hours of the day', or seasonal or unequal hours.

When the word 'hour' occurs in European literature through perhaps the fourteenth century, it commonly means a seasonal hour and not an hour as we know it. The hours were usually counted from sunrise or sunset, not from midnight and midday.

6. Celestial latitude and longitude

Celestial latitude and longitude are much like declination and right ascension except that they are based upon the plane of the ecliptic instead of the plane of the celestial equator. In order to find the celestial latitude and longitude of any point Q in the heavens, we start by passing the plane through Q that is normal to the ecliptic. The arc that this plane cuts on the celestial sphere from Q to the ecliptic is the celestial latitude β. The plane cuts the ecliptic in a point E, and the angle from the vernal equinox to E is the celestial longitude λ.

Celestial longitude is taken as positive if it is measured in an easterly direction from the vernal equinox. Celestial latitude is taken as positive for points that lie in a northerly direction from the ecliptic.

The transformation from α, δ to β, λ involves only a rotation about the direction of the equinoxes and is thus rather simple. It is

$$\left.\begin{aligned}\lambda &= \arctan\{(\cos\delta\sin\alpha\cos\epsilon+\sin\delta\sin\epsilon)/\cos\delta\cos\alpha\},\\ \beta &= \arcsin\{-\cos\delta\sin\alpha\sin\epsilon+\sin\delta\cos\epsilon\}.\end{aligned}\right\}\tag{6.1}$$

When the terms latitude and longitude are used, the context usually makes it clear whether the terrestrial or the celestial quantities are meant. Hence writers commonly use latitude and longitude without modifiers unless it is necessary to prevent ambiguity.

The classical planets (including the Sun and Moon) move nearly in the plane of the ecliptic. Further, the plane of the ecliptic changes its orientation in the heavens much less rapidly than the plane of the celestial equator. For these reasons, latitude and longitude are often more convenient coordinates than right ascension and declination. However, it is simpler and more precise to observe the latter pair of coordinates than it is to observe the former pair. Thus latitude and longitude are often used in theories of motion, but modern observations are almost always given in terms of right ascension and declination.

7. Planetary motions

I am still using the term planet to include the Sun and Moon. It is simplest to describe the motions of the Sun and Moon in geocentric terms and to describe the other motions in heliocentric terms. The descriptions, whether geocentric or heliocentric, are most simply referred to the plane of the ecliptic, with the vernal equinox as the basic reference line in that plane.

To high accuracy, the orbit of the Sun about the Earth is an ellipse, and this ellipse lies in the plane of the ecliptic. We need several quantities in order to describe the elliptic motion fully.

Two quantities specify the size and shape of the ellipse. In astronomy, the semi-major axis and the eccentricity are commonly used for this purpose. The semi-major axis is half the length of the long axis of the ellipse, while the eccentricity is related to the difference between the axes. The eccentricity is zero for a circle.

A third quantity specifies the orientation of the ellipse in the plane of the ecliptic. In order to specify the orientation, we usually start by finding the point of perigee, that is, the point at which the Sun comes closest to the Earth. We then give the angle Γ from the vernal equinox to the position of perigee. Γ is called the longitude of perigee.

While these three quantities tell us completely the path along which the Sun moves in space, they do not give enough information to let us calculate the point which the Sun occupies at a given time. Even when the path is known, calculation of position along the path is a complex matter that I shall not try to treat in this paper. I shall point out only that we need two additional quantities. One of them is the position at some conventional epoch. The conventional epoch used most often in current writing is noon on 1899 December 31, and the longitude at this time is called the longitude at the epoch.† The second additional quantity needed is the period of the motion, that is, the time needed to complete one circuit of the elliptical path. In the case of the solar orbit, the period is called the year.

To high accuracy, the orbit of the Moon about the Earth and the orbits of the other planets about the Sun are also ellipses. For each body, then, we must give the period, the position at some epoch, the orientation of perigee (in a heliocentric description, we use the term perihelion instead of perigee), the semi-major axis, and the eccentricity. However, we have an additional problem that we did not have with the solar orbit. The orbits of the Moon and of the other planets lie in planes that differ from each other and from the plane of the ecliptic. For each body, we must specify the plane of the orbit, and this takes two additional quantities. These quantities are illustrated in figure 6.

In figure 6, ♈ is the vernal equinox and the plane containing ♈ and N is the plane of the ecliptic. The plane NP is the plane of the orbit being considered. As the body moves along its orbit, it will clearly pass through the plane of the ecliptic at two points during each period.

† More accurately, the quantity usually given is the longitude of the mean Sun at the epoch. I shall not take the space to define the mean Sun here. If the reader makes a guess at what the mean Sun means, he will probably be close to being right.

These points are called the nodes of the orbit. At one node, the body passes from the north side of the ecliptic to the south side; this node is called the descending node. The other node, called the ascending node, is the point at which the body passes from the south side of the ecliptic to the north side. N represents the ascending node in figure 6.

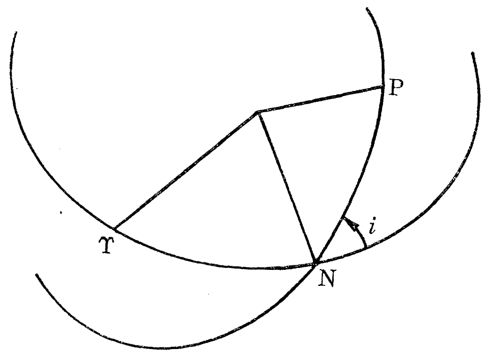

FIGURE 6. The quantities needed to specify the plane of an orbit for a body other than the Sun.

The angle from ♈ to N, measured in an eastward direction, is called the longitude of the ascending node, and it is frequently denoted by Ω. Ω is one of the two additional quantities needed to specify the plane of the orbit. The other is the inclination i, which is the dihedral angle between the planes, as shown in figure 6.

There are two methods in common use for giving the orientation of perigee (or perihelion). One is to give the angle ω measured in the direction of motion from N around to P in figure 6. This angle is called the argument of perigee or perihelion. The other method is to give a quantity Γ' defined by

$$\Gamma' = \Omega + \omega. \tag{7.1}$$

Γ' is often called the longitude of perigee or perihelion, although it is not the same as the celestial longitude of the point P. However, Γ' approaches the celestial longitude of P as the inclination i approaches 0.

Mercury has the most eccentric orbit of any major body in the solar system. Even for it, the axes of the orbit have the ratio 0.98; for other bodies, the ratio is even closer to 1. If a person saw the orbit of any major body drawn to scale, he would probably not recognize, with the unaided eye, the departure of the orbit from a circle. However, he would probably realize from the drawing that the centre of gravitation is not at the geometrical centre of the orbit.

The quantities that specify an elliptic orbit are often called the Kepler parameters of an orbit. If the orbits actually were ellipses, the Kepler parameters would be constants for a given planet. However, no real orbit is exactly an ellipse. The departures from elliptic motion can be described in a variety of ways. The way that is probably most valuable for intuitive thinking, as opposed to quantitative calculation, is to let the Kepler parameters vary with time.

We have already mentioned that the obliquity ϵ has been decreasing slowly throughout the historical period, and that the position of the vernal equinox is moving slowly westward through the fixed stars. The only other changes in parameters that need to concern us here are the changes in Ω and i for the lunar orbit.

The node of the lunar orbit rotates in a westerly direction around the plane of the ecliptic, making a complete revolution in about 18.61 years. This motion, and this time interval, are important in eclipse theory, as we shall discuss in the next section. This motion results mostly from the torque exerted on the Moon by the Earth's equatorial bulge. As such, it is a reaction to the precession of the equinoxes, which results mostly from the torque exerted on the bulge by the Moon's gravitation.

There are several types of variation in the lunar inclination. The largest one is produced by the Sun's gravitation, which causes an oscillation in the inclination. The amplitude of this oscillation is slightly less than 9′ of angle. The period of the oscillation is half of the period of revolution of the node, that is, it is about 9.3 years.

8. Eclipses

A celestial object B which is smaller than the Sun S casts a shadow whose geometry is shown in figure 7. First, there is the part shown solid between B and the point C; I shall call this part the inner umbra. Next there is the part shown solid but lying on the other side of C; I shall call this the outer umbra. Finally there is the part shown cross-hatched that surrounds the umbra; I shall call it the penumbra. We have an eclipse when another body enters, wholly or in part, either the umbra or the penumbra. It is clear that a body cannot enter the umbra without first entering the penumbra. Therefore the conditions for the occurrence of eclipses are governed by the penumbra.

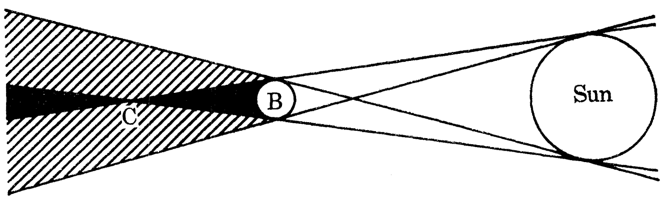

Figure 7. The geometry of the shadow cast by a body B. The part shown solid between B and C is the inner umbra, the part shown solid beyond C is the outer umbra, and the cross-hatched part is the penumbra.

When B represents the Moon, and the Earth enters the Moon's shadow, we have an eclipse of the Sun, as seen by people on Earth. People on the Moon, however, would see an eclipse of the Earth. When B represents the Earth, and the Moon enters the Earth's shadow, we have a lunar eclipse for Earth inhabitants and a solar eclipse for lunar ones. In the rest of the discussion, I shall consider only eclipses as seen by an inhabitant of the Earth.

Because the distances from the Sun to the Earth and moon are constantly changing, the dimensions of the shadows are continually changing. For our purposes, we can ignore this fact

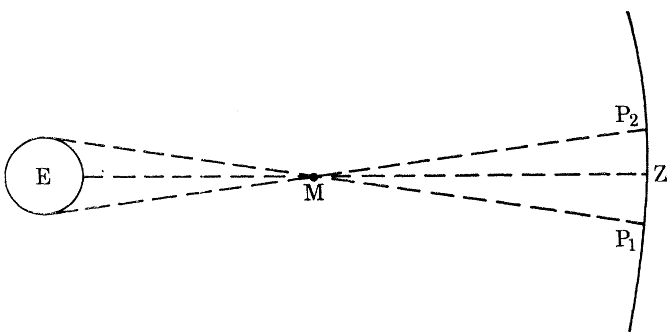

Figure 8. Positions of the centre M of the Moon, against the celestial sphere, for various observers on th e Earth E Any point between P_1 and P_2 is the apparent centre of the Moon for some observer on Earth.

and assume that the shadows have constant dimensions. We can also assume that the Sun has the same direction in the heavens regardless of the point on Earth from which it is viewed; that is, we can neglect the solar parallax. However, we cannot neglect the lunar parallax.

Consider where the centre of the Moon seems to be for an observer anywhere on Earth. Let M in figure 8 denote the centre of the lunar disk, and let E denote the Earth. For some observer, the Moon is in the zenith and its centre is at the point Z. There are many observers for whom the Moon is on the horizon. For one of these, the centre of the Moon is at the Point P_1 on the celestial sphere; for the diametrically opposite observer, the centre of the Moon is at the point P_2. Now draw the circle on the celestial sphere whose radius is P_1Z (which equals P_2Z) and whose centre is Z, and consider any point within this circle. There is some observer on Earth for whom the centre of the Moon lies at this point.

The radius $P_1Z = P_2Z = \pi_M$, say, is the quantity that was called the horizontal parallax in § 2.

Consider again the observer who sees the centre of the Moon at the point P_1. For him, the Moon blocks out a circle whose centre is P_1 and whose radius is s_M, say; s_M is called the semi-diameter of the Moon. For the diametrically opposite observer, the Moon blocks a circle whose centre is P_2 and whose radius is s_M. Clearly, if we draw a circle of radius $s_M + \pi_M$ centred on Z, any point within this circle is blocked from the view of some observer.

If any point of the solar disk comes within any part of the circle with radius $s_M + \pi_M$, some part of the solar disk will be obscured for some observer. Let the radius of the solar disk subtend an angle s_S, say, and let θ_{SM} denote the angular separation of the centres of the Sun and Moon. Then some part of the solar disk is obscured for some observer, that is, there is some sort of solar eclipse, whenever

$$\theta_{SM} < s_S + s_M + \pi_M, \tag{8.1}$$

approximately. The reader should remember that this relation and other relations to be obtained must be modified slightly to account for the solar parallax and other small effects.

In order to find a condition for lunar eclipses, let us imagine the Earth's shadow to be cast upon a sphere whose radius equals the distance to the Moon. Let the angular radius of the penumbral shadow be denoted by s_{ES}. Further, let U denote the centre of the shadow (this is the point on the celestial sphere diametrically opposite to the Sun), and let θ_{UM} be the angle from U to the centre of the Moon. Then there will be a lunar eclipse whenever

$$\theta_{UM} < s_M + s_{ES}. \tag{8.2}$$

The reader can readily convince himself, with the aid of a sketch, that $s_{ES} = \pi_M + s_S + \pi_S$, in which π_S is the solar parallax. If we neglect π_S, as we did in studying solar eclipses, we find from relation (8.2) that there is a lunar eclipse whenever

$$\theta_{UM} < s_S + s_M + \pi_M, \tag{8.3}$$

approximately.

The criteria in (8.1) and (8.3) are identical. That is, the numbers of solar eclipses and of lunar eclipses are equal to high accuracy. In spite of this, it is often said that there are considerably more solar eclipses. This statement results from neglecting a large class of lunar eclipses. However, the neglected eclipses are not readily visible to the naked eye. For observation with the unaided eye, there are more solar eclipses, if we consider all points on Earth. On the average, there are about 1.6 lunar eclipses and 2.4 solar eclipses per year that can be seen with the unaided eye. For a particular observer, however, there are many more lunar eclipses.

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF

The quantity that appears on the right of (8.1) and (8.3) is slightly less than 1.5°. Thus there can be an eclipse only when the centre of the Moon is within about 1.5° of the ecliptic. If the inclination of the lunar orbit were less than 1.5°, there would be a solar eclipse at each new Moon and a lunar eclipse at each full Moon. However, the inclination is about 5.1°, so that most new and full Moons occur without an eclipse.

If the Moon is to be close enough to the ecliptic to make an eclipse possible, it must be within about 17° of one of the points where it crosses the ecliptic; that is, the Moon must be within about 17° of one of its nodes at new or full Moon for an eclipse to happen. Suppose, for example, that one of the nodes is at longitude 0°, so that the other is at 180°, and suppose that the critical angle is exactly 17°. Then, in order to have an eclipse, the longitude of the Moon must either be between −17° and +17°, or it must be between 163° and 197°.

At new moon, the longitudes of the Moon and Sun are the same and at full Moon they differ by exactly 180°. Therefore the Sun must also lie within the ranges of longitude stated above in order for eclipses to occur. Suppose that the Sun has longitude 0° at the epoch E, say. Then it will be within one of the ranges of longitude from about E − 17 days to about E + 17 days. This interval of about 34 days is called an eclipse season. At all times within the season, the Sun is in a position that makes an eclipse possible. During one lunar revolution, the Moon takes on all longitudes. Since 34 days is greater than the orbital period of the Moon, the Moon is necessarily in the correct position for eclipses at some time during a season.

Thus there is at least one solar eclipse in each season, and there is at least one lunar eclipse if we consider all eclipses. If we consider only lunar eclipses that can be seen easily with the naked eye, however, it is possible for a season to pass without a visible lunar eclipse. In fact, it is possible for an entire year to pass without one.

If the nodes of the lunar orbit stayed in the same place, the eclipse seasons would come at the same time each year. However, as I said earlier, a node of the lunar orbit moves westward and makes a complete circuit in 18.61 years. This interval is therefore an important one in studying the recurrence of eclipses.

If the sun is at the position of one of the nodes at epoch E, it will be at the same node at the epoch E + 346.6 days, approximately. This interval of about 346.6 days is called an eclipse year; it is less than an ordinary year because of the westward movement of the node. There are two eclipse seasons in each eclipse year. On the average, there are slightly more than two seasons in an ordinary year. A given eclipse season moves through all the times of the ordinary year in an interval that averages about 18.61 years.

Concluding remarks

The purpose of this paper has been to supply some of the basic information about astronomy that is needed in order to study the role of astronomy in ancient civilizations. It is clearly not possible to supply all the information needed in a short paper, but I hope that this paper gives enough information to start the reader off on the subject.

The paper is odd in that it has given no references or citations. The reason is that the material in it has long since passed from the realm of research into the realm of standard knowledge. When so many references are available, to single out a few of them for citation would be invidious. Further, it is difficult to know the background of the reader in mathematics and the physical sciences, and thus it is difficult to know what references would be useful to him.

There is one reference, however, that should be mentioned. This is the *Explanatory supplement to The Astronomical Ephemeris and The American Ephemeris and Nautical Almanac* (London: Her Majesty's Stationery Office, 1961). This work explains how the tables in the ephemeris works of the United Kingdom and the United States are calculated, and it is invaluable for anyone who needs to do independent calculations about the behaviour of the solar system.

I thank my colleagues H. D. Black, B. B. Holland and R. E. Jenkins for valuable advice on the preparation of figures 1 to 5. I also thank Mr Jenkins for a critical review of the original typescript.

The preparation of this paper was supported by the Department of the Navy, under Contract N 00017-62-C-0604 with the Johns Hopkins University.

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF