JCAAC (2024), **1**, 5–14

JCAAC

METHODS & ALGORITHMS

Computational celestial navigation at sea

Luis Mederos[1]

[1]Consejo Superior de Investigaciones Científicas, Madrid, Spain. E-mail: lmederos@icmm.csic.es.

Keywords: astronomical navigation, problem of longitude

Abstract
Finding the geographic longitude of a ship when sailing in high seas, frequently referred to as *the longitude problem*, is one of the longest scientific problems in history that it took almost three centuries to be solved. The final solution to that problem, based on measuring the height of stars above the horizon, is briefly summarized in sections I and II. Section III is then devoted to explain how the traditional solution can be today implemented using simple calculations instead of the usual tedious graphical method.

Resumen
El cálculo de la longitud geográfica de un barco que navega en alta mar, denominado frecuentemente como *el problema de la longitud*, es uno de los problemas científicos más antiguos de la historia, que tardó casi tres siglos en resolverse. Además de resumir la solución final al problema, basada en la medida de la altura de las estrellas sobre el horizonte, presentamos en este artículo una versión de la solución tradicional basada en procedimientos simples, que sustituyen a los tediosos métodos gráficos usuales.

1. Introducción

Safe navigation requires sailors to know their position, i.e. latitude and longitude, quite often during the voyage. This was the case before the second half of the XV century, because the only navigation technique available was *coastal navigation*: sailors determined their position by simply taking references to known points on land. However, in the second half of the XV century, Portuguese sailors started ocean navigation around the south tip of Africa and the Indian ocean towards the spice islands and, soon after that, Castilian sailors tried to reach the same area sailing to the west across the Atlantic ocean, discovering America as a consequence. These ocean navigations spanned weeks, or even months, at sea without land references to determine the ship position, so that sailors had to turn their eyes to the sky in order to use stars as a reference to determine their position, giving rise to what we now know as *celestial navigation*.

Since Earth rotates from west to east, the position of a star in the sky, as seen by an observer at the Earth's surface, changes from east to west quite fast (approximately 15° every hour), but its position along the south-north direction is approximately constant. This implies that determining latitude by observing the stars is relatively easy compared to the problem of determining longitude. This is the reason why at the beginning of ocean navigation sailors already knew how to determine their latitude: They measured the height of the Sun above the horizon at its maximum, i. e. when the Sun is at the local meridian, a_m, and from the knowledge of the Sun's declination, δ, the latitude of the observer, φ, is immediately obtained, Fig. 1. Tables with the values of the Sun's declination were already available since the XIII

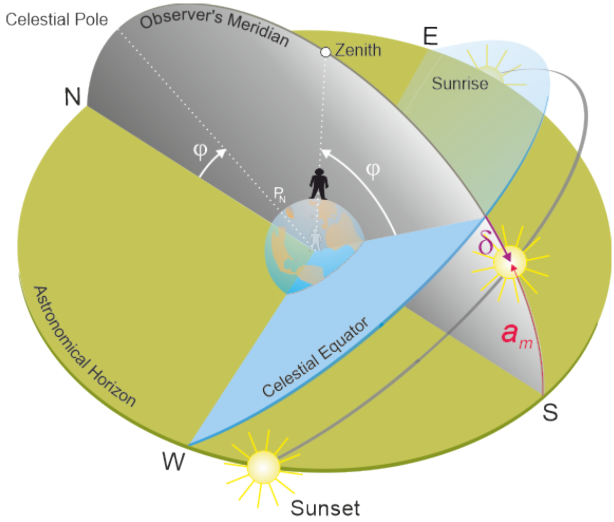

Figure 1. *Latitude from the meridian height of the Sun.*

century, when their construction was ordered by King Alfonso XIII. Later on, around 1478, a more complete publication, the *Almanaque Perpetuo*, was published by Abraham Zacuto in Salamanca. This publication contained declination values not only for the Sun, but also for the Moon and the planets Venus, Mars, Jupiter and Saturn.

A second alternative to determine latitude was also available at the beginning of ocean navigation, at least when sailing in the northern hemisphere. It was well known that the height of the celestial pole over the horizon is equal to the observer's latitude. Therefore, sailors could measure the height of the Polar star in the northern hemisphere, thus obtaining their latitude, Fig. 2.

The precision in the latitude was not high in those times: The instrument used by sailors to measure the height of the star above the horizon was the cross-staff, the much more precise sextant not being available until the XVIII century. Corrections to be applied to the measured height (refraction of light due to the Earth's atmosphere, horizon depression, parallax, etc.) were not known until that century. Moreover, the angular distance from the Polaris to the celestial pole $\sim 0.5°$ at present, but back in 1500 it was $3.5°$ due to precession of the Earth's rotation axis. Despite this problem, sailors managed to determine their latitude to within approximately $1°$, i. e. 60 nautical miles in the South-North direction, which was taken to be exact in those days [1].

And what about longitude? As already mentioned, determining the longitude from star observations is a much more difficult problem. This is because Earth's rotation causes the solution to the problem to rely on the precise measurement of time. Since the observer's longitude is the dihedral angle between the reference and observer's meridian planes, and a star moves approximately 15° to the west every hour (as seen by an observer on the surface of the Earth), the longitude could be determined by simply applying the idea already proposed as early as 1530 by Gemma Frisius: register the instant T_1 at which the star is on the reference meridian and, later, register the instant T_2 at which the star is on the observer's meridian. The longitude λ is nothing but the angle travelled to the west by the star during the time $T2 - T1$. This simple idea, however, gives rise to two problems that, in fact, underlay the impossibility to determine longitude in oceanic navigations until the middle of the XVIII century, a historical issue known as the *longitude problem* [**?**].

The first problem is that sailors in the middle of the sea could not observe the transit of the star at the

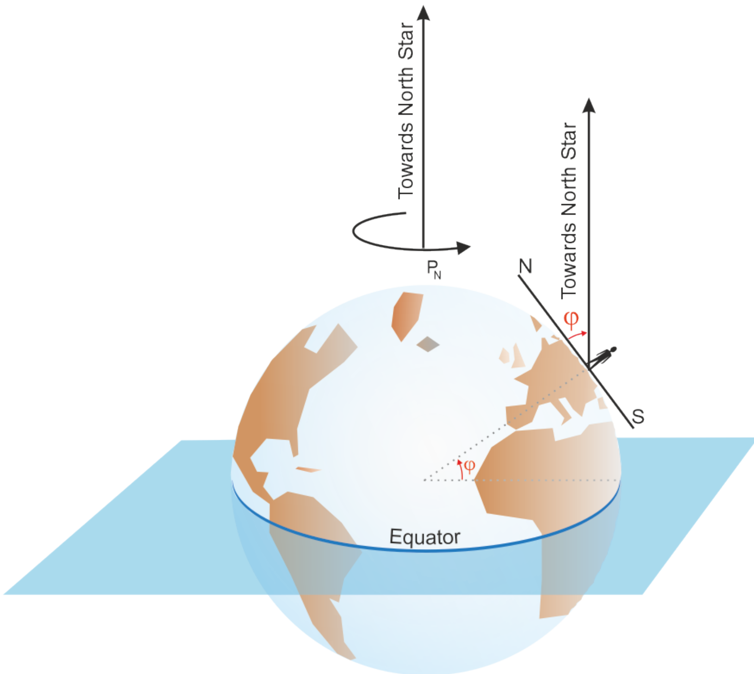

Figure 2. *Latitude from the height of Polaris.*

reference meridian, i.e. T_1 could not be measured. However, this problem had an easy solution since astronomers could predict the instants of those transits and published them in *nautical almanacs* carried on board. Obviously, the publications had to use the time scale of the reference meridian, i.e. Universal Time (UT), such that the same almanac was valid anywhere.

The second problem is that when the sailor observes the transit of the star across the local meridian, the registered instant of time T_2 has to be measured using the same time scale used by the almanac; only then will $T_2 - T_1$ be the time taken by the star to move for an angle λ. Sailors need to know the universal time at sea or, in other words, they have to *carry* UT on board. And since stars move to the west quite fast, roughly 15° every hour or 0.25′ every second, safe navigation require sailors to know UT within 1 second, Fig. 3. This was not possible until the second half of the XVIII century, thanks to the invention of the marine chronometer by John Harrison, about three centuries after the beginning of ocean navigation [6].

2. Celestial navigation

By the end of the XVIII century all the ingredients necessary to determine the boat position from star observations were already established:

- Astronomers could predict the precise positions of stars. *Celestial coordinates* of the stars were published in nautical almanacs, carried on board, as a function of the universal time UT. Use of these celestial coordinates allows to calculate the latitude and longitude of the star projections on the Earth surface, known as *geographic positions*, GP).
- Chronometers had been invented, providing the precise UT values of the star observations. Using the almanac, accurate values for latitude and longitude of the star projection on the Earth surface could be calculated.

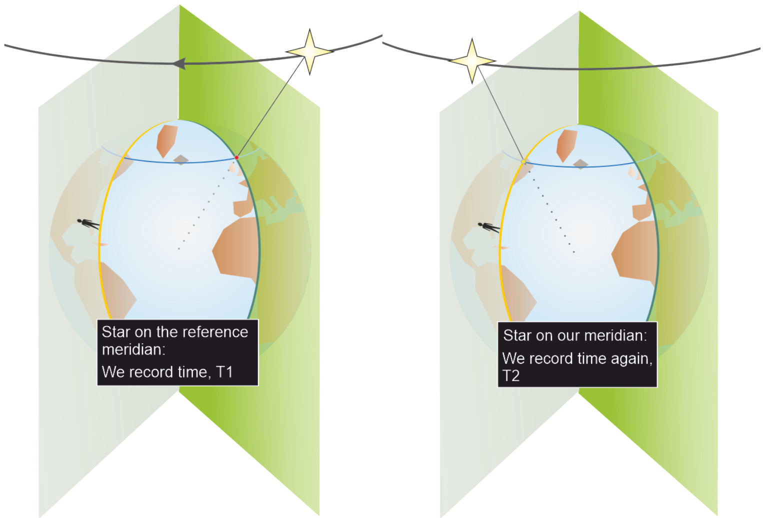

Figure 3. *Determining longitude from star observations necessarily requires the precise measurement of time.*

- Reflection instruments, such as the sextant, had already been developed, allowing the precise measurement of star heights on the horizon. Moreover, the corrections to be applied to the measured star height (refraction, parallax, etc.) to obtain the *true height*, a_v (i. e. the star height as seen from the centre of the Earth measured with respect to the *astronomical horizon* and without effects due to refraction, etc.) were known.

With all of these, we can easily determine our position at sea by observing the stars: We first measure the star height over the horizon using our sextant, taking care to write down also the precise UT time. We then apply all the needed corrections to obtain the true height a_v. Next, we use the registered UT to obtain, from the nautical almanac, the GP of the measured star. This point is the centre of a circle, drawn on the Earth surface. Our position will necessarily be located on this circle at the time of observation. Different observers located on that circle measure the same height in the same instant, Fig. 4; they simply need to look at different directions in order to see the star, i. e. they see the star with the same height but different azimuth. For this reason, this circle is frequently named *circle of equal altitude*.

Moreover, we also know the radius of the circle, which only depends on the height of the star on the horizon: the higher we observe the star, the nearer we are from the centre of the circle. The radius is simply $90° - a_v$, as it is obvious from Fig. 5.

In summary, by measuring a star height we can obtain a *line of position* over which we were necessarily located when we observed the star. By observing two stars, our actual position will be in one of the two points defined by the intersection of the two lines, Fig. 6. To discriminate between these two possibilities, we only need to take into account the approximate stars azimuths at the moment of their observation.

Obviously, computing the geographic coordinates of these intersection points was not possible until the second half of the XVIII century, when all the ingredients to obtain our position from stars observations were already available. The practical implementation of celestial navigation was based, and is based even today, on the fact that a sailor at sea always has a reasonably approximate position based on the situation of departure port, navigation course and speed. If this position is frequently updated, the sailor will always have an *estimated position*, S_e, which is close enough to the circle of equal altitude

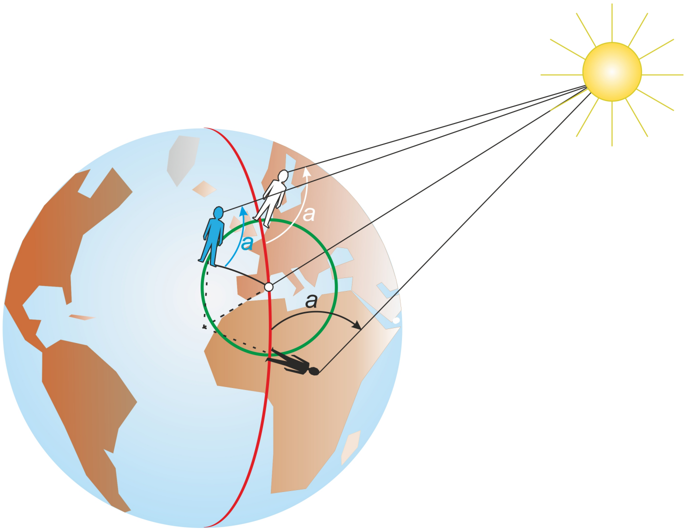

Figure 4. *The height of a star at a given time is the same when measured from any point of a circle centred on the star's geographic position.*

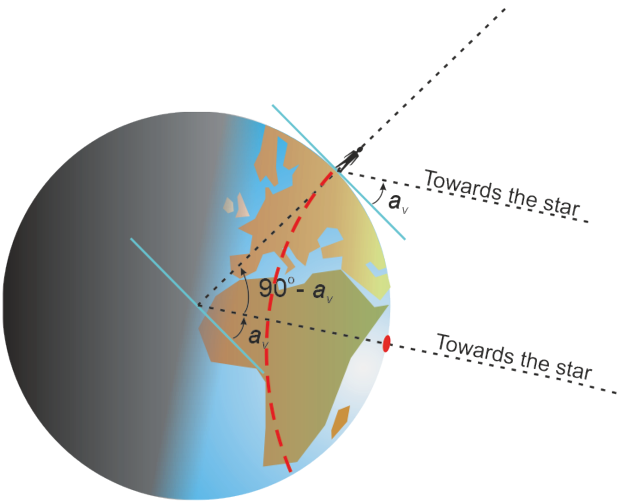

Figure 5. *The radius of the circle of equal altitude is simply* $90° - a_v$.

corresponding to an observed star and to the actual position at sea. Since the radius of the circle is enormous compared to the distance from our estimated position to the real position on the circle, it is reasonable to approximate the arc of circle at our position by a straight line that we can drawn on a Mercator chart. The problem of finding the intersection points coordinates was solved in practice graphically, by plotting the two arcs of equal altitudes near the estimated position, approximated by straight lines, on the chart and then looking for the coordinates of their intersection point [6].

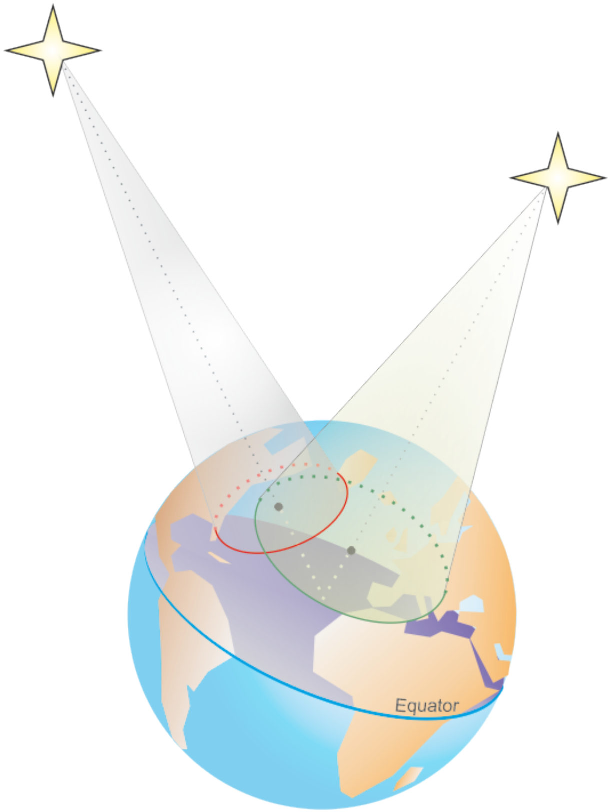

Figure 6. *By observing two stars, we can obtain our position on the Earth's surface.*

3. Computational Celestial Navigation

The traditional graphical solution to the problem of finding the coordinates of the intersection points of the two lines of position obtained from stars observations, summarized at the end of the previous section, depends on having a reasonably accurate estimated position. It also involves a tedious graphical work on the nautical chart. Today we can use a computer to avoid these two requirements.

3.1. Simultaneous observation of two stars

Let us start by discussing the simplest case where two stars are observed simultaneously. Of course, in this context 'simultaneously' means that there is a short time interval (a few minutes) between the two observations. Given the speed of our boat we can assume that our position has not changed during the observations. Then, at the end of our observations we will have two sets of values, one for each observed star,

$$(\mathrm{UT}_i, a_i, Z_i, \alpha_i, \delta_i)\,, \quad i = 1, 2, \tag{3.1}$$

where UT is the universal time instant when the star was observed, a is the true height of the star, i. e., the observed height corrected by depression of horizon, refraction, etc., Z is the observed approximate star azimuth, that will be only used to decide which of the two intersection points corresponds to our position, and α and δ are the right ascension and the declination of the star at the precise UT of measurement, as obtained from the nautical almanac on board.

We then consider a reference system with origin at the centre of the Earth, with the X axis pointing towards the Greenwich meridian, the Z axis pointing to the North Celestial Pole, and the Y axis following from the other to form a right-handed system. Let the two sidereal times corresponding to the observed UT_1 and UT_2 be θ_1 and θ_2. Then the unit position vectors $\hat{\boldsymbol{r}}_1$ and $\hat{\boldsymbol{r}}_2$ of the observed stars are given by

$$\begin{aligned}\hat{\boldsymbol{r}}_1 &= (x_1, y_1, z_1) = (\cos\delta_1\cos(\alpha_1-\theta_1),\ \cos\delta_1\sin(\alpha_1-\theta_1),\ \sin\delta_1)\\ \hat{\boldsymbol{r}}_2 &= (x_2, y_2, z_2) = (\cos\delta_2\cos(\alpha_2-\theta_2),\ \cos\delta_2\sin(\alpha_2-\theta_2),\ \sin\delta_2)\end{aligned} \tag{3.2}$$

Since the distance to the centre of the sphere is irrelevant (fortunately, ships are approximately at the same distance from the center of the Earth!) we solve the problem on a sphere of unit radius. Note that $\hat{\boldsymbol{r}}_1$ and $\hat{\boldsymbol{r}}_2$ point to the centres of the two circles of equal altitudes corresponding to the observed stars. The intersection points of these circles can be easily obtained by calculating the intersection (a straight line) of the two planes perpendicular to the directions $\hat{\boldsymbol{r}}_1$ y $\hat{\boldsymbol{r}}_2$, and then projecting the line on the unit sphere. Since we know the true heights of the stars, a_1 y a_2, the equations to solve are

$$\hat{\boldsymbol{r}}\cdot\hat{\boldsymbol{r}}_1 = \sin a_1, \quad \hat{\boldsymbol{r}}\cdot\hat{\boldsymbol{r}}_2 = \sin a_2, \quad |\hat{\boldsymbol{r}}|^2 = 1, \tag{3.3}$$

which will have two solutions, $\hat{\boldsymbol{r}}_+$ y $\hat{\boldsymbol{r}}_-$. Expressed in spherical coordinates, they will give the latitudes and longitudes of the two intersection points of the circles of equal altitude, i. e. our two possible positions. Since the distance between these two positions is enormous, we only need to take into account the approximate azimuths we have measured to decide on which of the two we were situated when the stars were observed.

The easiest way to solve equations (3.3) is to work in the XY plane to obtain the two planes intersection in parametric form, using z as parameter. Then the condition $|\hat{\boldsymbol{r}}|^2 = 1$ is imposed to obtain the two possible values z_+ and z_-, from which the other two coordinates x_+, x_- and y_+, y_- immediately follow. The two possible positions, $\hat{\boldsymbol{r}}_+$ y $\hat{\boldsymbol{r}}_-$ are then determined. The two first Eqns. (3.3) can be expressed as

$$\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sin a_1 - z z_1 \\ \sin a_2 - z z_2 \end{pmatrix} \tag{3.4}$$

Using z as parameter, the solution of this system of equations is

$$x = \frac{b_x + c_x z}{c_z}, \quad y = \frac{b_y + c_y z}{c_z}, \tag{3.5}$$

where we have introduced the following vectors:

$$\boldsymbol{c} \equiv \hat{\boldsymbol{r}}_1 \times \hat{\boldsymbol{r}}_2, \quad \boldsymbol{b} \equiv \boldsymbol{r}_2^\perp \sin a_1 - \boldsymbol{r}_1^\perp \sin a_2, \quad \boldsymbol{r}_i^\perp \equiv (y_i - x_i, 0),\ i = 1, 2. \tag{3.6}$$

If we now substitute x and y given by (3.5) into the condition $|\hat{\boldsymbol{r}}|^2 = 1$, we obtain a second-degree equation for z. The two solutions of this equation are the z coordinates of the two intersections of the circles of equal altitude:

$$z^2 + 2Bz + C = 0, \tag{3.7}$$

where

$$B = \frac{\vec{b}\cdot\vec{c}}{c^2} = \frac{b}{c}\cos\gamma, \quad C = \left(\frac{b}{c}\right)^2 - \left(\frac{c_z}{c}\right)^2. \tag{3.8}$$

The solutions to Eqn. (3.7) are:

$$x_\pm = \frac{b_x + c_x z_\pm}{c_z},\ y_\pm = \frac{b_y + c_y z_\pm}{c_z},\ z_\pm = \frac{b}{c}\left(-\cos\gamma \pm \sqrt{\left(\frac{c_z}{b}\right)^2 - \sin^2\gamma}\right). \tag{3.9}$$

Now we only need to express these two vectors in spherical coordinates to obtain the latitude and the longitude of the two intersection points, i. e. our two possible positions when we observed the stars.

3.2. Nonsimultaneous observation of two stars

In the case when there is a non-negligible time interval between the observations of the two stars (or between two observations of the same star) we cannot neglect the change of position from the instant of the first observation to the instant of the second observation. Let us assume that, in the interval between the two nonsimultaneous observations, our boat has sailed a distance D along the course over the sea R_v. Here we discuss a method to take this distance into account.

Let us suppose that, at the instant of our first observation, our position is the point on the circle of altitude represented by the ship in Fig. 7. As discussed before, since the radius of the circle is very large, the corresponding arc in the neighbourhood of our position can be approximated by a straight line. To use this line of position later, in combination with the one obtained from the second observation, we draw a parallel line displaced a distance D. Applying this argument to every point on the first circle of altitude, we arrive at the conclusion that ship displacement in the interval between the two observations can be accounted for by keeping the centre of the circle at the same point, but with a modified radius.
-0.3cm
Let a_1 be the true height corresponding to the first observation, the radius of the corresponding circle of altitude being is $90° - a_1$. The radius to be used will be $90° - a_1 \pm x$, with + if our boat is sailing away from the centre of the circle and − if the boat is sailing towards the centre, as can be seen in Fig. 7. Also, from the figure, it is easy to see that x is nothing but the displacement between the two observation along the direction defined by the azimuth, Z, of the star at the instant of the first observation. Since the distance D navigated between the two observations is considerably smaller than the radius of the circle, we can assume that the arc in Fig. 7 is flat in the neighbourhood of our ship and, consequently,

$$x = D\cos\alpha. \tag{3.10}$$

The angle α is obtained from the course of the ship, R_v, and the azimuth Z of the star during the first observation, which should in principle be measured as this is also useful to discriminate between the two intersection points. The problem is that measuring accurate azimuths from a sailing ship is not possible. Practical attempts to do it indicate that there is always an error of about $1 - 2°$. However, this error is not critical, as we are using the azimuth only to calculate the distance x sailed between the two observations, and this will always be very small compared to the radius of the circle of altitude. Therefore, using an approximate azimuth will not compromise our safety. Despite this, the next section discusses how to avoid the measurement of the star azimuth.

3.3. Avoiding measuring the azimuth

As mentioned at the end of the previous section, the star azimuth can only be measured approximately from a sailing boat. But in fact a virtually exact azimuth can be calculated, as will become clear shortly.

If we know our position, i. e. our latitude and longitude, we can easily calculate the exact azimuth of a star at a given instant of time (and also its height). Since the equatorial coordinates of the star are known from the nautical almanac, the corresponding horizontal coordinates, height and azimuth, can also be

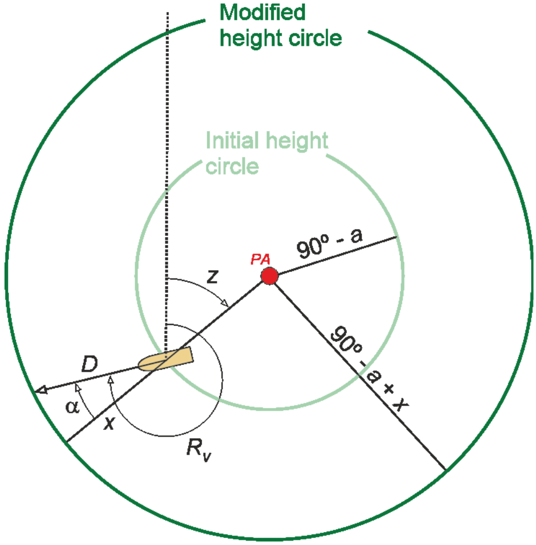

Figure 7. *Nonsimultaneous observations. To take into account the navigation between the two observations, we have to modify the radius of the circle of altitude.*

calculated [6]. However, out exact position is unknown. There are two ways to solve this problem.

The first one is to use our estimated position, S_e. As mentioned in Section 2, sailing requires keeping track of the course and speed over ground, so that the position is frequantly estimated and written down. This is, of course, an approximate position, because there will always exist some errors in course and speed. However, if things are done carefully, the estimated position will be sufficiently close, say with an error of a few nautical miles, to our (unknown) real position. Using the estimated position to calculate the star azimuth will produce a virtually exact result, because the centre of the circle of altitude is so far away (a few thousand nautical miles) that the direction towards we have to look at will be the same from two different points separated by only a few miles.

The second idea applies in the case where an estimated position is not available, i. e. when we have no idea about where we are. Since two nonsimultaneous observations are available, the intersection between the two circles of altitude can be calculated assuming that our boat has been at rest in the interval between the two observations. This will be an estimated position, instead of our real situation, since we have neglected the ship displacement. The estimated position will have an error of the order of the distance navigated in the interval between the two observations. So, we are back to the situation described in the previous paragraph and, therefore, we can now calculate the azimuth corresponding to the first observation. Moreover, if the boat speed is high and/or the time interval between observations is long, it is possible that the estimated position calculated neglecting the boat displacement between observations will be affected by an error which is too big to be neglected. In that case we only need to iterate the process.

Acknowledgments

I am very grateful to Dr. T. López Moratalla and Dr. E. Velasco for their stimulating comments over the years on the topic of this paper.

References

[1] Pedro de Medina, *Regimiento de navegación. Contiene las cosas que los pilotos han de saber para bien navegar, y los remedios y avisos quehan de tener para los peligros que navegando les pueden suceder*, 1563. In particular, concerning the determination of latitude, see *Libro Tercero*, titled *Del altura del Norte*. This book is available in digital form: *https://tinyurl.com/7y2wpuke*.

[2] William J. H. Andrewes (Ed.), *The quest for longitude*, Harvard University (1996).

[3] Daba Sobel, *Longitude: The true history of a lone genius who solved the greatest scientific problem of his time*, Harper Perennial (2005).

[4] See, for example, Luis Mederos, *Navegación Astronómica, 8ª Edición*, Tutor (2023), David Burch, *Celestial Navigation: A complete home study course*, Starpath (2015).

[5] Note that the two intersection points are usually separated by a very large distance, thousands of nautical miles, so the star azimuth will be significantly different when measured in the same instant from one of those points or from the other.

[6] See, for example, Luis Mederos, Navegación Astronómica, 8ª Edición. Editorial Tutor. Madrid (2023), Chapter 6.