Some New Dimensions in Sextant-Based Celestial Navigation Aspects of Position Solution Reliability with Multiple Sights

K.H. Zevering
Brisbane, Queensland, Australia

ABSTRACT: The traditional approach relying on sight reduction tables, a non-programmatic location of the position fix and an inadequate allowance for observation errors is still widely pursued and advocated. In the late 1970s the programmatic Least Squares method (LSQ) was introduced which determines a random error fix (Fix_Q) for any multiple sights combination. B.D Yallop & C.Y Hohenkerk (1985) expanded LSQ to incorporate the computation of the random error margin of a fix. Several marketed PDA-based programs apply LSQ, but none have fully incorporated the random error margin as a guide for the navigator. All existing LSQ applications have two drawbacks. One is, all observation error is attributed to random sources, whereas the possibility of systematic error has in fact a long theoretical and practical background in celestial navigation. Systematic error represents a bias in statistical random error theory and can and should be allowed for. A major drawback is that existing LSQ program applications incorporate the running fix technique (RFT) traditionally applied in coastal navigation. It has no general validity in celestial navigation. The position circle of an earlier celestial sight can only be mathematically correctly transferred when its Geometric Position (GP) is transferred for the run data. A final aspect of reliability is the strategy adopted at the sight planning stage. At least during twilight observations, navigators should aim at getting three or four sights with a total azimuth angle >180°, with three successive subsights on each body. In such configurations Fix_Q and Fix_S will be relatively close together, generally obviating the need to process the sights for possible systematic error.

1 THE TRADITIONAL VERSUS PROGRAMMATIC APPROACH

The calculator with trig functions makes the use of sight reduction (SR) tables redundant and avoids precession and nutation problems with AP 3270 Vol I. It does not overcome the limitations of the non-programmatic approach. Dedicated navigation calculators or PDAs would be a solution provided they run theoretically correct programs. Better known programs such as Starpilot, Astronavigation, Celestnav incorporate the traditional running fix technique (RFT) which is not generally applicable and cannot or cannot meaningfully compute the confidence ellipse (error margin) of the programmatic LSQ fix.

Complete power failure is often cited to support the use of SR tables and the non-programmatic approach. On today’s yachts such a contingency is for many reasons unrealistic and it would jeopardize nautical navigation in many ways. A better contingency plan is a Pocket PC running, say, Windows Mobile/Excel which can support the full programmatic approach in one or several spreadsheets for any number of sights.

The traditional approach relies on a number of precepts which are not borne out by modern methods and analysis:

1 The fix is presumed to lie somewhere in the middle or centre of a 'cocked hat' n-polygon (n ≥ 3).
2. The position of the vessel is put at the “greatest disadvantage”, but it means on a vertex or position line of the cocked hat n-polygon closest to a known danger. The error margin implied by the cocked hat area cannot be quantified.

3. Traditional RFT is used to account for the transfer of an earlier sight’s position circle for the run between sights. None of the above precepts have general validity. With i. we mean specifically that the fix is generally inside the cocked hat but in irregular n-polygons in an eccentric position (e.g. see Fig. 6). A fix with more than three sights may even lie outside the cocked hat n-polygon. One notion propounded by G. Huxtable and apparently taught in nav classes in England is that the chance of the celestial fix lying outside the 3-polygon is 75%, outside the 4-polygon 87 1/2% etc. The notion derives from random errors in compass bearings in coastal navigation (see Fig. 1)². With random + and - deviations, the terrestrial fix with three sights has a 75% chance of physically lying in peripheral cocked hat areas (Fig. 1, areas 1 to 6). There is only a 25% of the fix lying in cocked hat areas abc and def. The terrestrial construction is not possible unless (unknown) values are assigned to the deviations.

2. LSQ PECULIARITIES

The programmatic approach is based on LSQ (least squares method). The version demonstrated in B.D. Yallop and C.Y. Hohenkerk (1986) is applied to non-simultaneous sights and uses a RFT-based subroutine. We refer to this version as LSQ*. As RFT is not generally valid as a transfer technique, also LSQ* is not generally valid. For simultaneous sights the method is simply LSQ. Both programs compute a random error fix and its random error margin.

Yallop-Hohenkerk (1986, xx) assert that “it is possible to start from any position on the Earth, and, provided that \( \lambda_f \) (i.e. Long\(_{DR}\)) is kept in the range -180° to +180° and \( \phi_f \) (i.e. Lat\(_{DR}\)) in the range -90° to +90°, the position solution in most cases will begin to converge after a few iterations”. A position solution is indeed in principle independent from the initial assumption but different DR assumptions are necessary to force LSQ into finding the two alternative position fixes resulting from two sights with intersecting position circles. This may be demonstrated with a numerical example. For two simultaneous sights the two alternative position fixes can be found by applying either a double sight method like K-Z² or LSQ. Results with K-Z are:

\[
\begin{align*}
Q &= 0.387623; \quad R = 0.277141; \quad S = 1.196021; \\
Lat_1 &= 49.8408; \quad Lat_2 = -6.6652 \\
MD_{1,1} &= 79.7248; \quad MD_{1,1} = 15.3635; \quad MD_{2,1} = 65.7485; \quad MD_{2,1} = 29.3398 \\
Long_1 &= GHA_S + MD_{1,1} = 3.9715 W; \quad Long_1 = GHA_M - MD_{1,1} = 10.0048 E; \\
Lat_{1~2} &= 9.48408 N 3.9715 W (LHA_S 280.2752 \sim Zn_e 85.4518); \\
LHA_{1~2} &= 15.3635 \sim 205.3902; \quad Lat_{1~2} = 6.6652 S/10.0048 E (LHA_S 294.2515 \sim Zn_e 67.4722; LHA_M 29.3398 \sim Zn_M 307.5339). \end{align*}
\]

Note: SinLat\(_{1~2}\) = \((Q \pm R^2)/S; MD = meridian difference.

The same coordinates are found with LSQ using an appropriate initial position in each instance:

![Fig. 1. Errors in terrestrial compass bearings versus celestial observation errors](image)
3 THE GENERAL EQUATION OF A POSITION CIRCLE

A position circle is represented on a Mercator chart by the general equation. To plot a relevant segment of it on the chart it is necessary to know the coordinates of its GP (GHA and Dec) and its radius (90° - Ho). To transfer the position circle of an earlier sight for a run it is necessary to determine the transferred GP in terms of its new coordinates.

To quote the ANM on this point: "If the observer is in a ship and there is a run between sights, the first position circle must be transferred for the run.

This can be done by transferring the geographical position and then drawing the circle."^5

The relevant segments of the position circles and their point(s) of intersection can only be drawn on a (small-scale) chart with a pair of compasses if their altitudes are large and radii (zenith distances) consequently small. In the general case when zenith distances are very large such a construction must use the general equation, which was circumvented in the pre-electronic era by applying traditional RFT known from coastal navigation.

4 THE TRANSFER ISSUE

With traditional RFT the transferred position line is supposed to represent a position circle segment. An article of faith for traditional RFT supporters is that the transferred position 'backward-projected' for the run data is postulated to lie on the original position circle. This concept derives directly from coastal navigation. Thus in Fig. 4a, the position J1 backward-projected from Fix'Q is supposed to lie on the original position circle of the Moon.

This is nonetheless an untenable proposition for two main arguments. One is that coastal or terrestrial RFT on which celestial RFT is patterned is itself in principle an application of GD-UT transfer (Fig. 3).

By nominating longitude intervals for x the corresponding latitude values for \( \phi \) can be computed and plotted as in Fig. 2 for two sights where the GHA, Dec and Ho are for convenience simplified as 0°, 0° and 30°, 1°, 45° respectively. There is a run due north of 1° between these sights, so that the transferred GP becomes GHA* = 0°, Dec* = 1°. In Fig. 2, the longitude scale is fixed but the latitude scale is derived from the meridional parts formula. The programmatic fix in this case may be obtained without plotting by applying K-Z.

The other argument is that celestial RFT cannot specify the transferred position circle in terms of its radius 90° - Ho and relocated GP.

4.1 The terrestrial transfer analogy

Points A and B in Fig. 3 represent landmarks with known height and also the GPs of two celestial bodies with large altitudes. Assume in the terrestrial case the distance to A and to B is determined by vertical sextant angle; no bearings are used.

The relevant position fix on the chart is at F where the transferred position circle with radius \( r_A \) from A' (PC*1) and the last sight's position circle with radius \( r_B \) from B (PC2) intersect. Had only the
bearings been used the fix would also have been in F, assuming the celestial azimuths could be observed accurately. In the celestial case when \( r_A \) and \( r_B \) are short zenith distances, \( DF \) virtually equals \( AA' \) and the angle at \( D \) the course (\( \alpha \)). In the terrestrial case a running fix as at \( F \) can therefore be found in all instances by applying the GD-UT transfer principle if \( r_A \) and \( r_B \) are known. In the celestial case, \( F \) is also found in all instances with GD -UT by transferring the 1st sight’s GP for the run data and plotting the relevant sections of \( PC^*_1 \) and \( PC_2 \) (with the general equation).

The solution for the SH is forced by assuming a DR at say 59.4000 S/85.1000 E. The results with LSQ* are:

<table>
<thead>
<tr>
<th></th>
<th>Northern Hemisphere:</th>
<th>Southern Hemisphere:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moon LHA</td>
<td>353.3483</td>
<td>92.8268</td>
</tr>
<tr>
<td>Moon X</td>
<td>-0.9481</td>
<td>-0.2310</td>
</tr>
<tr>
<td>Moon Y</td>
<td>0.1074</td>
<td>-0.9258</td>
</tr>
<tr>
<td>Moon TanA</td>
<td>-0.1132</td>
<td>-0.7654</td>
</tr>
<tr>
<td>Moon A</td>
<td>-6.4609</td>
<td>75.9896</td>
</tr>
<tr>
<td>Moon Z</td>
<td>173.5391</td>
<td>37.4288</td>
</tr>
<tr>
<td>Moon e</td>
<td>0.0000</td>
<td>322.5712</td>
</tr>
<tr>
<td>DR (Sun)</td>
<td>50.3496 N/14.0474 W</td>
<td>59.4000 S/85.1000 E</td>
</tr>
<tr>
<td>Fix’Q (4th it)</td>
<td>50.49178 N/13.8418 W</td>
<td>59.0949 S/85.8293 E</td>
</tr>
</tbody>
</table>

When GD-UT+K-Z is applied, the results are:

<table>
<thead>
<tr>
<th></th>
<th>Northern Hemisphere:</th>
<th>Southern Hemisphere:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moon LHA</td>
<td>353.6209</td>
<td>92.5608</td>
</tr>
<tr>
<td>Moon X</td>
<td>295.3460</td>
<td>34.2859</td>
</tr>
<tr>
<td>Moon Y</td>
<td>106.7762</td>
<td>323.3593</td>
</tr>
<tr>
<td>Moon TanA</td>
<td>173.7959</td>
<td>256.3396</td>
</tr>
<tr>
<td>Moon A</td>
<td>104.0104</td>
<td>323.3593</td>
</tr>
<tr>
<td>Moon Z</td>
<td>255.9896</td>
<td>323.3593</td>
</tr>
<tr>
<td>Moon e</td>
<td>0.0000</td>
<td>323.3593</td>
</tr>
<tr>
<td>DR (Sun)</td>
<td>50.5117 N/13.8323 W</td>
<td>59.3735 W/85.1076 E</td>
</tr>
<tr>
<td>FixQ</td>
<td>50.5117 N/13.8323 W</td>
<td>59.3735 W/85.1076 E</td>
</tr>
</tbody>
</table>

The plot for the NH is shown in Fig 4a. The distance between Fix’Q and FixQ for the NH is 1’.25, but for the SH it is 27’.75. With RFT/LSQ*, the backward-projected positions respectively from Fix1 (NH) and from Fix2 (SH) are postulated to lie on the original position circle at J1 (NH) and J2 (SH) (see sketch Fig. 4b). If this were a correct proposition, the position circle passing through J1 and J2 should have a zenith distance equal to 90° - Ho = 72°.5883.

But the zenith distance of the position circle passing through these two points can only be 72°.8551°.

This apparent contraction in zenith distance is not the only problem because none of the great-circle segments in Fig. 4b indicated with broken lines or the angles they make can be evaluated. Also, the Zn-values of the Moon (earlier sight) shown above for the final LSQ* iteration are the azimuth bearings from the backward-projected positions J1 and J2 and not from Fix1 and Fix2.

LSQ* cannot compute the actual azimuth bearing from the fix on the transferred GP. Even if this were possible, the properties (coordinates of the GP and zenith distance) of the earlier sight’s position circle transferred for the run data cannot be computed. We call this the ‘Achilles heel’ of LSQ*. The only correct way to transfer the GP of an earlier sight is GD-UT.

4.2 The specification of the position circle transferred with RFT/LSQ*

When zenith distances are large the plane-geometric properties of the terrestrial analogy no longer apply and it is impossible to correctly specify a transferred position circle like \( PC^*_1 \) in Fig. 3 as the locus of points like \( J^* \) and \( F \) whose backward-projected positions are postulated to lie on \( PC_1 \) at \( J \) and \( D \). This can be demonstrated in different ways but here we choose the Moon-Run-Sun example found in the ANM³ (see Fig. 4).

---

![Fig. 3. The GD-UT principle of terrestrial RFT](image)
Fig. 4a. Plot of LSQ* transfer of an earlier sight (The Moon-Run-Sun case in the ANM)

Fig. 4b. Sketch of the implications of the LSQ* transfer as in Fig. 4a

5 RANDOM AND SYSTEMATIC ERRORS

Error intercepts (e) indicate a possible combination of random and systematic error. Both types of error will displace a position line parallel to itself in a certain direction along its Zn axis. Systematic error ($e_s$) is caused by instrument error (e.g. I.E, Dip) and is equal in magnitude and sign for each sight in the collection. LSQ computes a random error fix ($Fix_Q$) and systematic error statistically constitutes bias. Maximum permissible systematic error ($e_s$ max) can be removed as a correction to the altitudes. LSQ will then determine a fix ($Fix_S$) corrected for $e_s$ max. This approach in fact reconciles conventional LSQ concerned with random error and traditional bisector constructions which allow for maximum systematic error in consistent 3-polygons.

For practical navigational reasons and also lack of space we will confine the discussion to 3- and 4-polygons. The main factor affecting the difference in location (distance $d$) between $Fix_Q$ and $Fix_S$ is 'total azimuth angle' (TAZ), which is the smallest angle or arc enclosing all azimuth (Zn) bearings. A statistic $\delta$ expressing the difference in location is the ratio of $d$ to average error intercept distance: $\delta = d / \Sigma |e_i|/n$. The consistency of a collection of sights is seen from the signs of the e’s found with LSQ: it is consistent if they all have the same sign; if one or more signs are different it is inconsistent (also see Fig. 1).

Navigators are to take at least three subsights on the same body for screening aberrant sightings in the usual manner. The programmatic approach makes it possible to process all included subsights independently, rather than the averages of the subsights’ GMT and $H_s$ data. The various ramifications of such an approach cannot be discussed here.

During the short twilight periods navigators should concentrate on the best star triads indicated in AP 3270 Vol 1 which fall in group II. Four sights should be planned to fall in groups IV and V. As configurations in group V are preponderant, the effect of possible systematic error on the location of the LSQ fix is in this way minimized. Collections I, II and III can be predicted from the approximate azimuths of the sights known beforehand, but this is not possible for IV and V. When three sights fall in group I, a fourth sight may be taken so that the overall combination will have the favourable azimuth distribution (i.e TAZ > 180°). Best results are generally obtained from two intersecting pairs in which the sights in each pair are widely separated in azimuth.

<table>
<thead>
<tr>
<th>Properties</th>
<th>3-polygon:</th>
<th>4-polygon:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TAZ &lt; 180°</td>
<td>TAZ &gt; 180°</td>
</tr>
<tr>
<td></td>
<td>Inconsistent I</td>
<td>Consistent II</td>
</tr>
<tr>
<td>1</td>
<td>% chance</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td>$Fix_Q$</td>
<td>inside</td>
</tr>
<tr>
<td>3</td>
<td>$Fix_S$</td>
<td>outside</td>
</tr>
<tr>
<td>4</td>
<td>$\delta$</td>
<td>large</td>
</tr>
<tr>
<td>5</td>
<td>$\delta$ range</td>
<td>1.5 to 3.0</td>
</tr>
</tbody>
</table>

Notes: Internal vertex angles (IVA) of an n-polygon are < 180°. Item 1 indicates the chance of getting certain combinations in twilight observations; if 4-polygons with TAZ > 180° are planned, the binomial distribution is as shown.
5.1 Procedures for obtaining $e_S$ max

Sketched in Fig. 5 are vertices 1 and 2 of an n-polygon and adjacent IVAs A and B. The perpendicular $h$ on side AB is computed with the known values of the two IVAs and the known distance (d) between the two vertices. The distance $d$ equals $d'\text{Lat}/\cos(Zn \pm 90)$, where $d'\text{Lat}$ derives from the vertex latitudes computed with K-Z. The perpendicular $h$ indicates $|e_S\text{ max}|$ if it is the smallest among similar perpendiculars dropped on the other sides: $H'=H_0 \pm h_{\text{min}}/60$, where $H'_0$ is a sight’s altitude corrected for $e_S\text{ max}$; the sign of $h_{\text{min}}$ is opposite to the sign of the corresponding error intercept (e). An IVA is consistent (C) when the error intercepts of its sides have the same sign; otherwise it is inconsistent (IC). Two adjacent IVAs may form one of the following sequences: CC, IC-IC, C-IC (a and b) and IC-C (a and b). The IVA angles are computed as:

$$C: |Zn_A-Zn_B| < 180 \rightarrow IVA = 180 - |Zn_A-Zn_B|;$$
$$|Zn_A-Zn_B| > 180 \rightarrow IVA = |Zn_A-Zn_B| - 180,$$

$$IC: |Zn_A-Zn_B| < 180 \rightarrow IVA = |Zn_A-Zn_B|;$$
$$|Zn_A-Zn_B| > 180 \rightarrow IVA = 360 - |Zn_A-Zn_B|.$$

Needed in the computation of $h$ are the $\frac{1}{2}$IVAs; if IVA = IC, $\frac{1}{2}$IVA = $\frac{1}{2}(180 - IVA)$. If the angles of two adjacent IVAs are indicated as respectively A and B, then A' and B' indicate their supplements. The formulas for computing the plane-geometric perpendicular $h$ for different IVA sequences are shown in Fig. 5.

The approximate method is sufficient and has the advantage that the distance $d$ need not be computed. From the results for the Sun1-Sun2-Sun3 case shown in Fig 6 $|Zn_1-Zn_3| > 180$, thus $IVA_{\text{apex}} = 251.7313 - 180 = 71.7313$; $e_1 = 0'.75$ and $e_3 = 0'.86$. These values substituted in the approximate method give approximate $|e_S\text{ max}| = 1'.36$.

5.2 Application of $e_S$ max to three and four sights

The navigator is in practice faced with two situations. One is the need for updating the DR position for which FixQ is generally sufficient. The other is the need for considering the vessel’s position in the presence of a known danger. Depending on the consistency of the sights, possible systematic error may significantly affect the location of the fix but not necessarily the error margin. A third situation is simply that the navigator is prepared to speculate on the preponderance of either type of observation error and also to consider error margins at less than 95% probability.

Removing $e_S\text{ max}$ from a consistent 3-polygon (group II) will virtually eliminate all residual random error and FixS is at the point of intersection of the bisectors. As FixQ and FixS are relatively close, in both practical situations mentioned FixQ and its error margin will suffice. Space does not permit to demonstrate this, but with four inconsistent sights (group V) the two fixes will be very close and their respective error margins remain practically the same. Again, FixQ will suffice in all circumstances.

In Fig 6. are worked examples of an inconsistent 3-polygon (group I) and consistent 4-polygon (group IV). The inconsistent 3-polygon is avoidable in twilight observations but running fixes on the Sun remain most important and will often form an inconsistent collection. The chance of getting four consistent star sights, let alone larger collections of stars as dished up in contrived examples in the literature is remote. In the three Sun sights case, the DR position is wildly out and indicates that the danger has been cleared. If FixQ and its error margin is accepted, the vessel might not change course.
Fig. 6. Cases falling respectively in Group I and group IV.
FixS with three inconsistent sights always lies outside the 3-polygon. It is prudent in this case to accept FixS and change course.

Assuming long runs between sights as in this case is common in the literature but a habit which should not be emulated in practice. It compounds the (unknown) inaccuracies of the DR record. As the run data are used to transfer the position circles of earlier sights, a large discrepancy between the fix and the run record as in this case also tends to invalidate the fix. In other respects the DR position should be completely ignored as irrelevant. LSQ will determine the same fix regardless of a wide range of assumed (DR) positions.

In the consistent 4-polygon case, noted is first the eccentric location of the fix. The distance between the two fixes is significant. Allowing for $e_\text{S}$ max substantially reduces the error margin, but in the presence of a known danger, the prudent navigator would adopt FixQ.

REFERENCES

Admiralty Navigation Manual, 1938, Volumes II and III.
Bennett G.G., Celestial Navigator, 2003-2007;
Blewitt M., Celestial Navigation for Yachtsmen, 1975, 1996;
Cunliffe T., Celestial Navigation, 2001;

---

1 See G. Huxtable, quoted in K.H. Zevering – The Navigator’s Newsletter (Foundation for the Promotion of the Art of Navigation), no. 88, p. 11-12.
3 Data from M. Blewitt (1975, p. 30 and p. 33). GHA, Dec and $H_o$ of Sun 284.2467, 18.4050 and 20.5150; Moon 19.3350, 15.4900 and 53.4550.
5 ANM, Vol. II, p. 43.
7 With GD-UT, the coordinates of the GP may be transferred with the rhumbline equations and are indicated as GHA* and Dec*. With the vessel’s movement, the GP of the sight taken at A transfers to A’ (see Fig. 3). AA’ is strictly not a rhumbline but part of a great circle. The angle of cut of this great circle with the meridians through respectively A and A’ is not constant as the mercatorial bearing $\alpha$ at A suggests. Applying the convergency (c), for $\alpha$ could be substituted $\alpha' = \alpha \pm \frac{1}{2}|c|$, where $c = (d/60)\sin\tan(\text{Mean Dec})$. In $\Delta{PA'A'}$, Dec* of the transferred GP at A’ may be determined with $\text{Lat}_A$, $\alpha'$ and $d/60$. Angle (arc) APA’ may be determined with LatA, Dec* and $d/60$, from which follows GHA*. It can be shown that even with very large displacements these adjustments have a negligible effect compared to the rhumbline definitions of Dec* and GHA*.
9 ANM, Vol II, p 191-195; only the fix for the NH is worked out, with cosine-haversine and the traverse table. The GMT data for Moon and Sun are: 05 59 45 and 08 40 10. The data for GHA, Dec and $H_o$ are: Moon 7.9783, -22.0400, 17.4117; Sun 309.1783, 5.2050, 19.9450. The assumed DR position in the NH (Moon sight) is 50.1667 N/50.8333 W; course 70°, speed 12 kn/hr.
10 $\cos\beta_1 = (\sin H_o - \sin Lat_1 \sin Dec) / \cos Lat_1 \cos Dec$ and $\cos\beta_2 = (\sin H_o - \sin Lat_2 \sin Dec) / \cos Lat_2 \cos Dec$. The angles $\beta_1$ and $\beta_2$ are known. $\sin Ho = \frac{c1c2(a1+a2)+b1c2+b2c1}{(c1+c2)}$, where $a1 = \cos\beta_1$; $a2 = \cos\beta_2$; $b1 = \sin Lat_1 \sin Dec$; $c1 = \cos Lat_1 \cos Dec$; $b2 = \sin Lat_2 \sin Dec$; $c2 = \cos Lat_2 \cos Dec$. $\sin Ho = \frac{c1c2(a1+a2)+b1c2+b2c1}{(c1+c2)}$.
11 Data from G.G. Bennett (2003, p. 164).
12 Data from M. Blewitt (1975, p 39)
13 E.g see M. Blewitt (1975, 1997); T. Cunliffe (2001).