

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 $>180^\circ$, 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 1. 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 ($n \geq 3$).

- 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. Huxtable¹ 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 87½% 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.

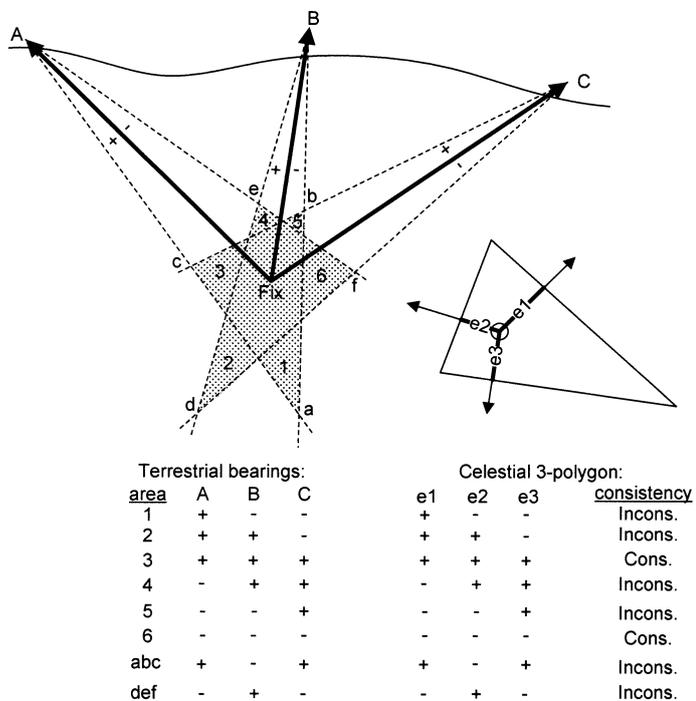


Fig. 1. Errors in terrestrial compass bearings versus celestial observation errors

The + and - deviations in terrestrial bearings are equated with the chance of getting *toward* (+) and *away* (-) intercept (p) combinations in celestial

navigation. The idea pursued by Huxtable is that the celestial fix will only lie inside the triangle when the p-values are ++ + or - - -, which is simply incorrect. The p-values are only corrections for inaccurate DR position. With three sights the random error fix will always lie inside the triangle and the observation errors e1, e2 and e3 are measurable. There is a theoretical 25% chance of getting a consistent 3-polygon (e_i = +++ , or ---). Based on the azimuth distributions of stars listed in AP 3270 Vol. 1 this chance is more like 40%.

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 λ_f (i.e Long_{DR}) is kept in the range -180° to $+180^\circ$ and ϕ_f (i.e Lat_{DR}) in the range -90° to $+90^\circ$, 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 example³. For two simultaneous sights the two alternative position fixes can be found by applying either a double sight method like K-Z⁴ or LSQ. Results with K-Z are:

$$Q = 0.387623; R = 0.277141; S = 1.196021;$$

$$\text{Lat}_1 = 49.8408; \text{Lat}_2 = -6.6652$$

$$\text{MD1,1a} = 79.7248; \text{MD1,1b} = 15.3635; \text{MD2,1a} = 65.7485; \text{MD2,1b} = 29.3398$$

$$\begin{aligned} \text{Long}_1 &= \text{GHA}_S + \text{MD1,1a} = \text{GHA}_M - \text{MD1,1b} = 3.9715 \text{ W}; \\ \text{Long}_2 &= \text{GHA}_S + \text{MD2,1a} = \text{GHA}_M - \text{MD2,1b} = 10.0048 \text{ E}; \\ \text{Lat}_1 \sim \text{Long}_1 &= 49.8408 \text{ N} / 3.9715 \text{ W (LHA}_S \text{ 280.2752} \sim \text{Zn}_S \text{ 85.4518}; \end{aligned}$$

$$\begin{aligned} \text{LHA}_M \text{ 15.3635} &\sim 205.3902); \text{Lat}_2 \sim \text{Long}_2 = 6.6652 \\ \text{S} / 10.0048 \text{ E (LHA}_S \text{ 294.2515} &\sim \text{Zn}_S \text{ 67.4722; LHA}_M \text{ 29.3398} \sim \\ \text{Zn}_M \text{ 307.5339).} \end{aligned}$$

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

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

	Northern Hemisph.:		Southern Hemisph.:	
DR	49° 50'.0 N/4° 20'.0 W		6° 42'.0 S/10° 30'.0 E	
	1 st iteration		1 st iteration	
	Sun	Moon	Sun	Moon
Zn	85.1796	204.8322	67.3309	307.1141
p	13'.99	-6'.35	-26'.44	24'.73
	3 rd iteration		3 rd iteration	
Zn	85.4518	205.3902	67.4722	307.5339
e	-0'.00	0'.00	-0'.00	0'.00
Fix _Q	49.8408 N/3.9715 W		6.6652 S/10.0048 E	

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.

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 ($90^\circ - H_0$). 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.

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 φ can be computed and plotted as in Fig. 2 for two sights where the

GHA, Dec and H₀ are for convenience simplified as 0°, 0°, 30° and 45°, 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 $90^\circ - H_0$ and relocated GP.

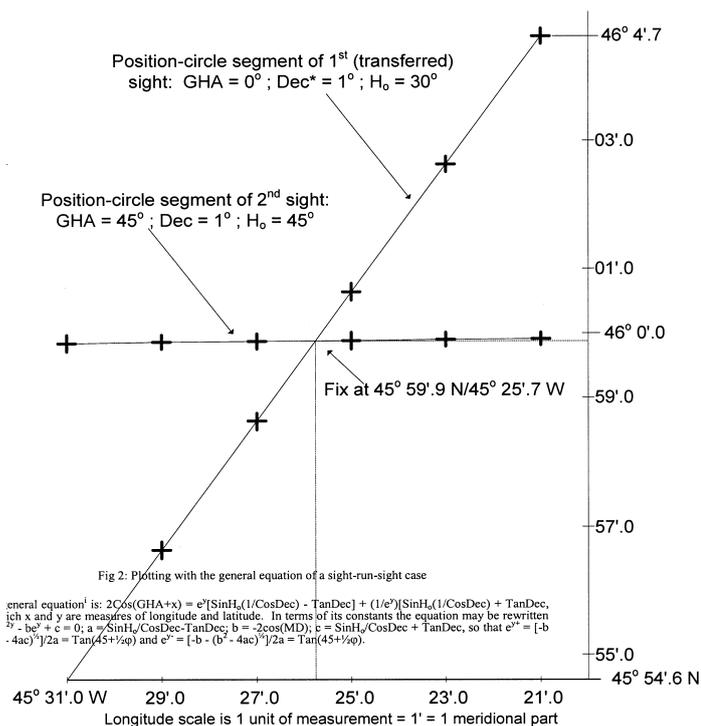


Fig. 2. Plotting with the general equation of a sight-run-sight case

The general equation⁶ is: $2\text{Cos}(GHA+x) = e^y[\text{Sin}H_0(1/\text{CosDec}) - \text{TanDec}] + (1/e^y)[\text{Sin}H_0(1/\text{CosDec}) + \text{TanDec}]$, in which x and y are measures of longitude and latitude. In terms of its constants the equation may be rewritten as: $ae^{2y} - be^y + c = 0$; $a = \text{Sin}H_0/\text{CosDec} - \text{TanDec}$; $b = -2\text{cos}(MD)$; $c = \text{Sin}H_0/\text{CosDec} + \text{TanDec}$, so that $e^{y+} = [-b + (b^2 - 4ac)^{1/2}]/2a = \text{Tan}(45 + 1/2\phi)$ and $e^{y-} = [-b - (b^2 - 4ac)^{1/2}]/2a = \text{Tan}(45 - 1/2\phi)$.

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 (PC_2) intersect. Had only the

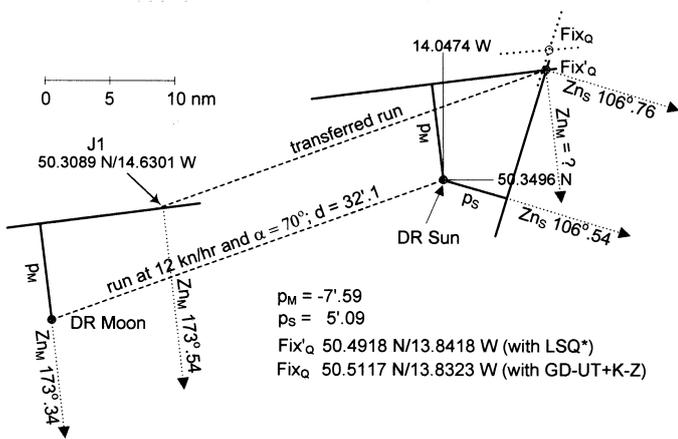


Fig. 4a. Plot of LSQ* transfer of an earlier sight (The Moon-Run-Sun case in the ANM)

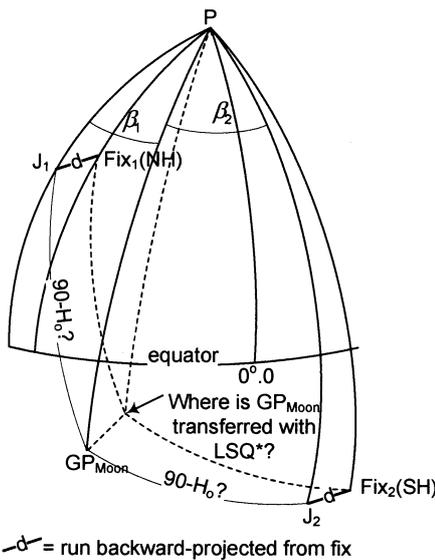


Fig. 4b. Sketch of the implications of the LSQ* transfer as in Fig. 4a

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 Z_n 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 (Z_n) bearings. A statistic \check{s} expressing the difference in location is the ratio of d to average error intercept distance: $\check{s} = d \div \Sigma|e_i|/n$. 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. $TAZ > 180^\circ$). Best results are generally obtained from two intersecting pairs in which the sights in each pair are widely separated in azimuth.

Properties		3-polygon:		4-polygon:		
		TAZ < 180°	TAZ > 180°	TAZ < 180°	TAZ > 180°	
		Inconsistent I	Consistent II	Inconsistent III	Consistent IV	Inconsistent V
1	% chance	60	40	33	12? (P)	87? (P)
2	Fix_Q	inside	inside	in- or outside	inside	in- or outside
3	Fix_S	outside	inside	in- or outside	inside	in- or outside
4	\check{s}	large	small	large	small	smallest to negligible
5	\check{s} range	1.5 to 3.0	around 0.2	0.6 to 2.7	around 0.6	0.0 to 0.16

Notes: Internal vertex angles (IVA) of an n-polygon are $< 180^\circ$. Item 1 indicates the chance of getting certain combinations in twilight observations; if 4-polygons with $TAZ > 180^\circ$ 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 $|d'Lat|/\cos(Z_n \pm 90)$, where $d'Lat$ derives from the vertex latitudes computed with K-Z. The perpendicular h indicates $|e_s \max|$ if it is the smallest among similar perpendiculars dropped on the other sides: $H'_o = H_o \pm h_{\min}/60$, where H'_o is a sight's altitude corrected for $e_s \max$; the sign of h_{\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:

$$C : |Z_{nA} - Z_{nB}| < 180 \rightarrow IVA = 180 - |Z_{nA} - Z_{nB}|;$$

$$|Z_{nA} - Z_{nB}| > 180 \rightarrow IVA = |Z_{nA} - Z_{nB}| - 180, \text{ and}$$

$$IC : |Z_{nA} - Z_{nB}| < 180 \rightarrow IVA = |Z_{nA} - Z_{nB}|;$$

$$|Z_{nA} - Z_{nB}| > 180 \rightarrow IVA = 360 - |Z_{nA} - Z_{nB}|.$$

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.

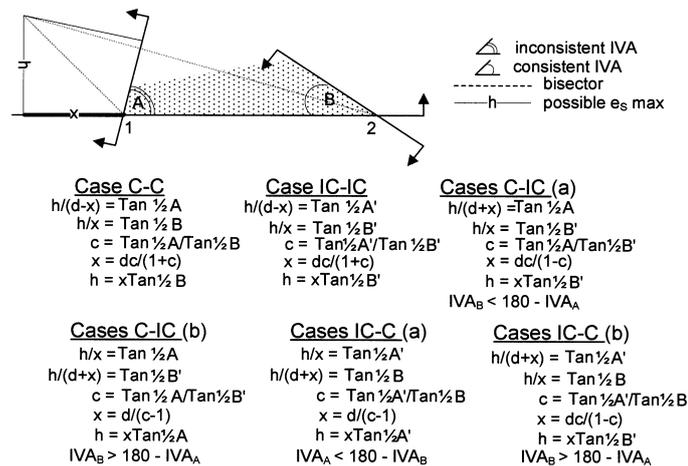


Fig. 5. Determining the h -values for consistent cocked hat 3-polygons and for n -polygons ($n \geq 3$)

With group I sights, the perpendicular h from the intersection of the bisector of the consistent (apex) IVA and this IVA's opposite side determines $e_s \max$:

Approximate method

$$\begin{aligned} h/(d-x) &= \tan \frac{1}{2}A = \tan \frac{1}{2}IVA_{\text{apex}} & \sin(\frac{1}{2}IVA_{\text{apex}}) &\approx \frac{1}{2}(|e_3| + |e_1|)/z \\ h/x &= \tan \frac{1}{2}B' = \tan IVA_{\text{base}} & z &\approx \frac{1}{2}(|e_3| + |e_1|)/\sin(\frac{1}{2}IVA_{\text{apex}}) \\ c &= \tan \frac{1}{2}IVA_{\text{apex}} / \tan IVA_{\text{base}} & \frac{1}{2}(|e_3| + |e_1|)/h &\approx z/(z + |e_2|) \\ x &= dc / (1+c) & \text{where } |e_2| &\text{always } < (\text{bisector} - z) \\ h &= x \tan IVA_{\text{base}} & h &= \text{approx. } e_s \max \approx \frac{1}{2}(|e_3| + |e_1|)(z + |e_2|)/z \\ \text{Use smallest } IVA_{\text{base}} & & e_1 \text{ and } e_3 &\text{ are the error intercepts to the } IVA_{\text{apex}} \text{ sides} \end{aligned}$$

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 $|Z_{n1} - Z_{n3}| > 180$, 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 $|e_s \max| = 1'.36$.

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 Fix_Q 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 $e_s \max$ from a consistent 3-polygon (group II) will virtually eliminate all residual random error and Fix_S is at the point of intersection of the bisectors. As Fix_Q and Fix_S are relatively close, in both practical situations mentioned Fix_Q 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, Fix_Q will suffice in all circumstances.

In Fig 6. are worked examples of an inconsistent 3-polygon (group D)¹¹ and consistent 4-polygon (group IV)¹². 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 literature¹³ is remote. In the three Sun sights case, the DR position is wildly out and indicates that the danger has been cleared. If Fix_Q and its error margin is accepted, the vessel might not change course.

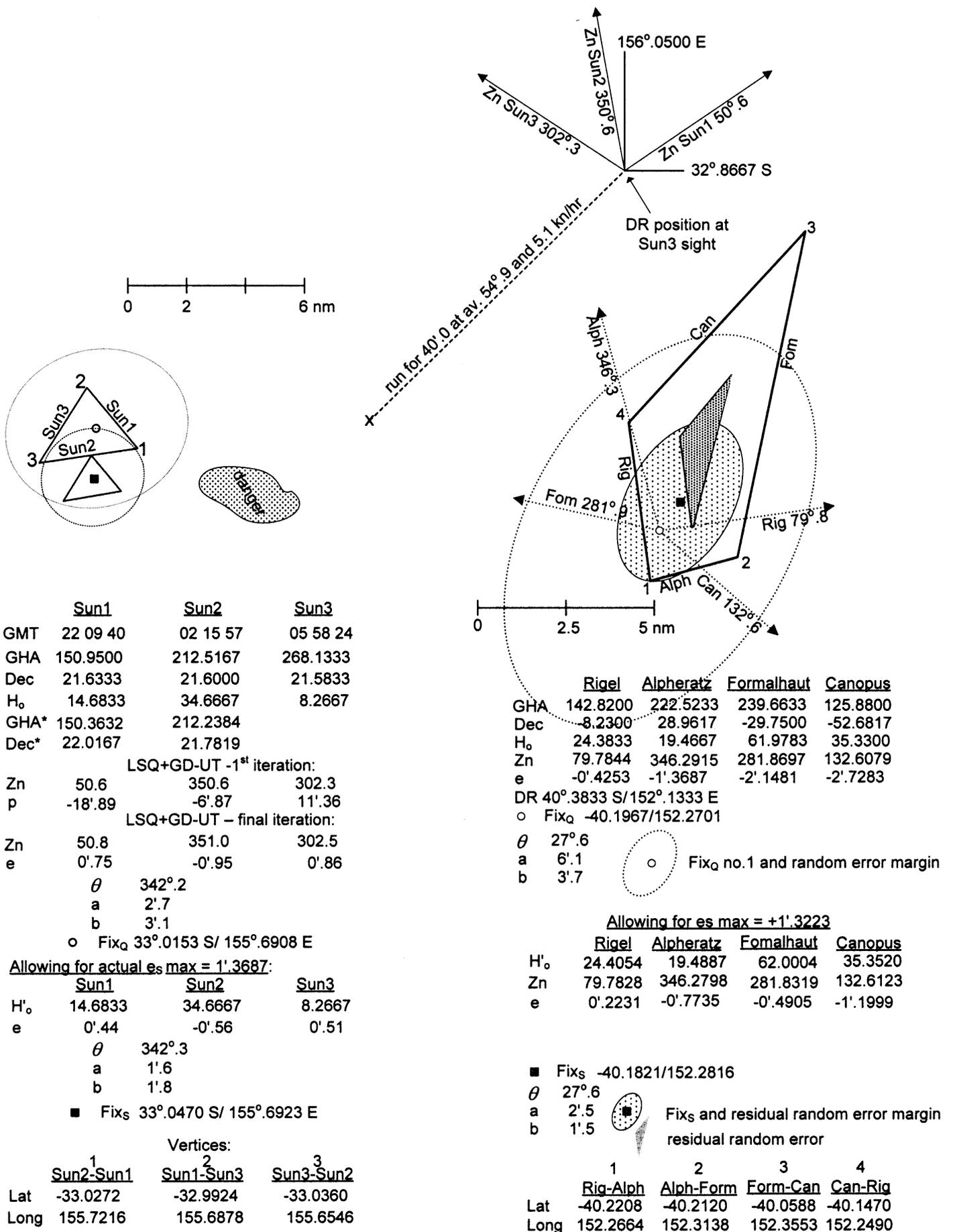


Fig. 6. Cases falling respectively in Group I and group IV

Fix_S with three inconsistent sights always lies outside the 3-polygon. It is prudent in this case to accept Fix_S 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_S max substantially reduces the error margin, but in the presence of a known danger, the prudent navigator would adopt Fix_Q.

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;
 Yallop B.D. & Hohenkerk C.Y., *Compact Data for Navigations and Astronomy*, 1986-1990.

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.

² See ANM, Vol. III, p. 165-166.

³ Data from M. Blewitt (1975, p. 30 and p. 33). GHA, Dec and H₀ of Sun 284.2467, 18.4050 and 20.5150; Moon 19.3350, 15.4900 and 53.4550.

⁴ For the K-Z algebraic double sight position solution method see K. Herman Zevering – “The K-Z Position Solution For The Double Sight”, *European Journal of Navigation*, Vol. 1, No. 3 (and 4), 2003, p. 43-46.

⁵ ANM, Vol. II, p. 43.

⁶ See ANM, Vol. III, p. 36, 39-40.

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 α at A suggests. Applying *the convergency* (c), for α could be substituted $\alpha' = \alpha \pm \frac{1}{2}|c|$, where $c = (d/60)\text{Sin}\alpha\text{Tan}(\text{MeanDec})$. In $\Delta PAA'$, Dec* of the transferred GP at A' may be determined with Lat_A, α' and d/60. Angle (arc) APA' may be determined with Lat_A, 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*.

⁸ See the discussion in Forum, *The Journal of Navigation* (2006), 59, p. 521-529.

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 H₀ 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 $\text{Cos}\beta_1 = (\text{Sin}H_0 - \text{SinLat}_{j1}\text{SinDec})/\text{CosLat}_{j1}\text{CosDec}$ and $\text{Cos}\beta_2 = (\text{Sin}H_0 - \text{SinLat}_{j2}\text{SinDec})/\text{CosLat}_{j2}\text{CosDec}$. The angles β_1 and β_2 are known. $\text{Sin}H_0 = [c_1c_2(a_1+a_2)+b_1c_2+b_2c_1]/(c_1+c_2)$, where $a_1 = \text{Cos}\beta_1$; $a_2 = \text{Cos}\beta_2$; $b_1 = \text{Sin Lat}_{j1}\text{SinDec}$; $c_1 = \text{CosLat}_{j1}\text{CosDec}$; $b_2 = \text{Sin Lat}_{j2}\text{SinDec}$; $c_2 = \text{CosLat}_{j2}\text{CosDec}$. $\text{Sin} H_0 = [c_1c_2(a_1+a_2)+b_1c_2+b_2c_1]/(c_1+c_2)$.

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).