UNLIMITED

AD-A272 937

TR 93004

DEFENCE RESEARCH AGENCY

Technical Report 93004

April 1993

A New Procedure for Orbit Determination Based on Three Lines of Sight (Angles Only)

by

R. H. Gooding

93-28106

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

DTIC
ELECTE
NOV 19 1993
S B D

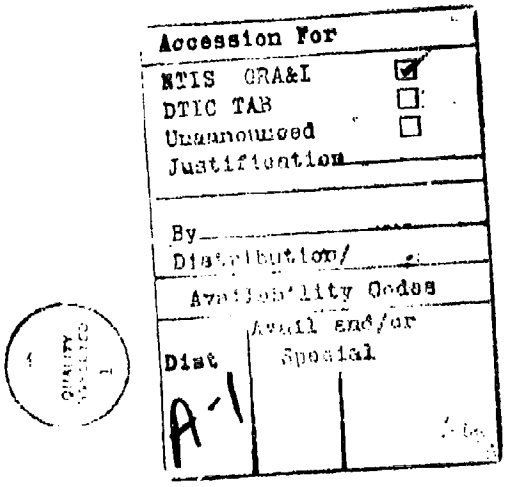
Farnborough, Hampshire

93 11 16 045

UNLIMITED

Requests for reproduction should be made to:

Intellectual Property Department
Defence Research Agency
R69 Building
Farnborough, Hampshire GU14 6TD



DEFENCE RESEARCH AGENCY
Farnborough

Technical Report 93004

Received for printing 13 April 1993

A NEW PROCEDURE FOR ORBIT DETERMINATION BASED ON THREE LINES OF SIGHT (ANGLES ONLY)

by

R. H. Gooding

SUMMARY

A new procedure has been developed for the general solution of the minimal angles-only problem in which an orbit is determined from three line-of-sight observations. The basis of the approach is a higher-order Newton correction of the assumed values for two of the unknown ranges, appeal being made to the author's (published) universal solution of Lambert's orbital boundary-value problem.

The new procedure is free of the inherent limitations of the traditional methods of Laplace and Gauss, these methods being outlined in a summary of previous approaches to this classical problem. In particular, the observations are permitted to span several revolutions when the orbit is elliptic; multiple solutions can be obtained; and there is no restriction on the configuration of the three observing sites.

The procedure has been carefully tested, some of the examples being taken from the literature. A number of test problems have been solved that would have failed by existing methods.

Departmental Reference: Space 689

© Crown Copyright (1993)
Defence Research Agency
Farnborough Hampshire GU14 6TD UK

LIST OF CONTENTS

1 INTRODUCTION

Determining a body's orbit from three line-of-sight observations is one of the classical problems of orbit dynamics, being significantly more difficult than solving either of Kepler's equations[1,2] or Lambert's problem[3]. It resembles Lambert's problem (the derivation of a Keplerian orbit from positions at two given times) in its minimal nature, the difference being that we have two items of data (angles only) at three times instead of three items of data (range also) at two times. The greater complexity arises because details of the observers' locations must be available in the angles-only problem, whereas they are entirely irrelevant in Lambert's problem since absolute positions are known. In pre-radar days it was the angles-only problem alone that was of real practical significance; the two may now be regarded as of equal importance.

To present the problem in its starkest form, we suppose that nothing is known about the hypothetical orbit, other than through the three observations, except that the observed body can be assumed to have a Keplerian motion about a given force centre (parent body), C, of gravitational strength μ; it is even conceivable that the observations are not really of the same body, in which case the problem would be an illusory one. Thus all possible solutions are sought, the problem being usually described as one of 'initial' (or 'preliminary') orbit determination. If an approximate orbit can be established from the minimal data available, then a more accurate orbit can be derived later, using as many observations as desired, but that would be an entirely different problem (see, for example, Refs 4 and 5 for details of the author's computer program, PROP), in which the techniques of linearization (relative to the initial orbit), least squares and differential correction can be applied. (These techniques, and a program such as PROP, are irrelevant until a sufficiently accurate set of orbital parameters is available for differential correction to converge.)

To specify the angles-only problem precisely, we suppose that observations are available at times t_j ($j = 1,2,3$) from the sites O_j defined by the vectors $\underset{\sim}{R}_j$, and that they are given (after conversion from pairs of angles if necessary) by the unit vectors (direction cosines) $\underset{\sim}{\lambda}_j$. An underlying (inertial) coordinate system, $Cxyz$, is assumed to have been adopted, relative to which the known locations $\underset{\sim}{R}_j$ may be specified as (X_j,Y_j,Z_j). The unknown orbital positions, P_j, are referred to by the vectors $\underset{\sim}{r}_j$, with components (x_j,y_j,z_j) and the unknown ranges* are denoted by ρ_j, so that the equation for the geometry at time t_j may be expressed as

$$\rho_j \underset{\sim}{\lambda}_j = \underset{\sim}{\rho}_j = \underset{\sim}{r}_j - \underset{\sim}{R}_j \quad . \tag{1}$$

---

* Since the 'directions', as opposed to 'anti-directions', of the observations are known, only solutions with all $\rho_j > 0$ are valid; however, an 'invalid' solution becomes valid, and a 'maximal family' of solutions is generated, if the problem is transformed by moving the O_j far enough 'backwards' along the sight-lines.

It is normal, but not obligatory, to take the t_j as such that $t_1 < t_2 < t_3$, and it is convenient to use the notation t_{jk} for the differences $t_k - t_j$. Since most approaches to the problem are iterative but with fewer than three iteration variables, the observations are not normally all dealt with in the same way; in particular, one pair of suitable iteration variables consists of ρ_1 and ρ_3, and another of ρ_2 and $\dot{\rho}_2$. In view of the implied special role for $j = 2$ (which will be assumed even when t_2 is **not** the intermediate time), it will be convenient to assume (for r and ρ in particular) that, when the suffix is suppressed, $j = 2$.

The importance and antiquity of the angles-only problem are reflected by its appearance in standard text-books – Refs 6 to 11 are cited (in chronological order) as providing good background material. For specialized papers with new approaches, and in the context of geocentric satellite orbits rather than heliocentric comets or minor planets, Refs 12 and 13 are to be recommended. In the present paper, section 2 provides summaries of the classical solution procedures devised by Laplace and Gauss some two centuries ago. The methods of these giants are limited by certain assumptions, in particular the assumption that the observations span only a limited orbital arc, and this is why more general methods, suitable for the Space age, have had to be devised. The approach of these more general methods, also summarized in section 2, is to iterate on the ranges associated with two of the observations[12], or even all three[13], whilst making the orbit computations in essentially closed form (*ie* without truncation).

The function of the present Report is to give a detailed description of the author's own new method for the angles-only problem. It is a range-iteration method, but unlike previous ones in that it places an efficient and universal algorithm[3] for the subordinate Lambert problem at the heart of the solution procedure. The description of the method appears in section 3, and section 4 shows how the method, as implemented in a Fortran-77 computer program, performs with some example problems from the literature; a number of special examples have also been constructed, and show that the new method can solve problems that existing methods could not.

To conclude this introduction, it is noted that in the operational use of an angles-only solution procedure it is essential to include 'aberration' corrections for the finite speed at which light- or radio-waves travel down the sight-lines. The way to make these corrections is covered by the text-books, and they have not been included in the procedure described in sections 3 and 4. It has already been indicated that orbit perturbations, also, will be ignored.

2 PREVIOUS APPROACHES

There have been two traditional approaches to the solution of the angles-only problem. They are known as Laplace's method and Gauss's method, though the latter

(which is more flexible) is also associated, in the evolution of its various versions, with the names of Lagrange and Gibbs. The basic principles of the two methods are summarized in sections 2.1 and 2.2, and then section 2.3 provides an introduction to the more recent range-iteration methods.

2.1 Laplace's method

Laplace's method suffers from two major limitations; these have not, however, been a serious drawback in the traditional application to the determination of initial orbits for comets and minor planets. The first limitation is the necessity for all three observations to be, in essence (but see the second remark at the end of this summary), from the same location (with due allowance for the site's rotation with the Earth, and/or its heliocentric motion, as appropriate). The second limitation is to a set of observations that only span a relatively short arc of the orbit. The reason for these limitations is that the method is directed towards the determination of ρ_2 and $\dot{\rho}_2$ (after which, $\underset{\sim}{r}_2$ and $\dot{\underset{\sim}{r}}_2$ are known, and hence also the orbit) by following an implicit assumption: namely, that the $\underset{\sim}{\rho}_j$ are instantaneous values of a well-defined continuous vector, $\underset{\sim}{\rho}$, for which first- and second-order time derivatives are also well-defined.

The effect of the foregoing assumption, and the two basic limitations which underlie it, is that the three $\underset{\sim}{\lambda}_j$ may be transformed (via numerical-difference formulae) into the triad $\underset{\sim}{\lambda}$ $(= \underset{\sim}{\lambda}_2)$, $\dot{\underset{\sim}{\lambda}}$ and $\ddot{\underset{\sim}{\lambda}}$. This leaves ρ $(= \rho_2)$, $\dot{\rho}$ and $\ddot{\rho}$ as three unknown scalar quantities, to be connected via the vector dynamical condition that must hold at time $t_{(2)}$, viz

$$\ddot{\underset{\sim}{r}} = -\mu \underset{\sim}{r}/r^3 . \tag{2}$$

The vectors $\underset{\sim}{r}$ and $\ddot{\underset{\sim}{r}}$, appearing linearly in equation (2), can immediately be eliminated via equation (1), with $\ddot{\underset{\sim}{R}}$ assumed known, but there remains a fourth unknown scalar, r. Hence we need a fourth scalar equation to supplement the three corresponding to the vector equation derived from (1) and (2). The fourth equation is provided by the geometry, since the cosine formula, applied to the triangle COP with the angle at O denoted by ψ, or (equivalently) the inner product of $\underset{\sim}{r}$ with itself via (1), gives

$$r^2 = R^2 + \rho^2 - 2R\rho \cos \psi , \tag{3}$$

where $\cos \psi = -\hat{\underset{\sim}{R}} \cdot \underset{\sim}{\lambda}$ (using the usual notation for unit vectors) and so is known.

Elimination of $\ddot{\rho}$ and $\dot{\rho}$ (via scalar multiplication by $\underset{\sim}{\lambda} \wedge \dot{\underset{\sim}{\lambda}}$) from the vector equation leads to an approximate linear relation between ρ and $1/r^3$. Then the elimination of either r or (preferably) ρ, by use of (3), leads to an octic equation in the other quantity. This equation is analysed in the literature[7], but celestial mechanicians have found it more

fruitful to work with an equivalent trigonometric equation in the angle ϕ at P – both r and ρ can be expressed in terms of ϕ by two applications of the sine formula to the triangle COP. The form of the trigonometric equation* is

$$\sin^4\phi = M \sin(\phi + m) , \qquad (4)$$

where M (> 0) and m are quantities known[6] in terms of $\underset{\sim}{R}$, $\underset{\sim}{\lambda}$, $\dot{\underset{\sim}{\lambda}}$, $\ddot{\underset{\sim}{\lambda}}$ and the subsidiary ψ.

The fundamental equation, whether an octic polynomial in r or having the form (4) in ϕ, has[7] at most three relevant roots (such that $r > 0$ and, correspondingly, $0 < \phi < \pi$) as well as (in general) exactly one irrelevant one (for which $r < 0$ and $\phi > \pi$); of the 'relevant' roots, only those are valid for which $\rho > 0$ and, correspondingly, $0 < \phi < \pi - \psi$. In the traditional application, however, for heliocentric orbits observed from the surface of the Earth, the radius of the Earth is essentially negligible and the observer's location is itself effectively in an unperturbed heliocentric orbit; thus there must exist a trivial solution (giving the observer's orbit) with ρ_j essentially zero for all j. So there are traditionally at most two non-trivial solutions. If two solutions exist and one cannot be rejected on the basis that it involves $\rho_j < 0$ for some j or (alternatively) a set of impossible orbital elements, then the normal way to distinguish the true solution from the spurious one would be by the introduction of a fourth observation.

In the familiar Earth-satellite application, on the other hand, there is no trivial solution, since a ground-based observer (even on the equator) cannot be in geocentric orbit. Thus there may be as many as three *a priori* valid non-trivial solutions, with the consequent need to reject up to two. If the 'observer' is satellite-based, however, then we are back to the existence of a trivial solution with at most one other to be rejected.

Various techniques have been applied to the solution of the fundamental equation, usually involving iterative refinement by the Newton-Raphson process or a similar technique. It must be remembered, however, that the equation can provide no more than an approximation to the solution for ρ_2 and r_2 (whence $\dot{\rho}_2$ and therefore $\dot{r}_2$, since $\dot{R}_2$ is assumed known), as the transformation from $\underset{\sim}{\lambda}_1$ and $\underset{\sim}{\lambda}_3$ to $\dot{\underset{\sim}{\lambda}}_2$ and $\ddot{\underset{\sim}{\lambda}}_2$ is itself only approximate. Thus Laplace's method, in its simplest form, leads merely to initial estimates for ρ and $\dot{\rho}$. If an accurate solution of the problem is required, an iterative differential-correction procedure (on ρ and $\dot{\rho}$) can then be used, to fit directly to (residuals in) $\underset{\sim}{\lambda}_1$ and $\underset{\sim}{\lambda}_3$. Alternatively (but less accurately), and more in the spirit of Laplace's method[6], the initial solution can be improved by using it to estimate values for the hitherto neglected components of higher derivatives, $\dddot{\underset{\sim}{\lambda}}$ etc; in this approach, the accuracy remains limited by the use of assumed values for $\dot{\underset{\sim}{R}}$ and $\ddot{\underset{\sim}{R}}$.

---

* The trigonometric approach was due to Gauss – an octic equation appears in Gauss's method that is of the same form as the octic equation in Laplace's method, as we shall see.

This summary of Laplace's method is concluded by three remarks. The first is that the inherent errors in the method are least when t_2 really is intermediate between t_1 and t_3, the optimum situation being when it is just half way. Secondly, with reference to the limitation to observations from the same location, it is noted by Escobal[10] that, when this does not hold, conceptual values for $\dot{\underset{\sim}{R}}$ and $\ddot{\underset{\sim}{R}}$ may nonetheless be taken from the same numerical-difference formulae as are used for $\dot{\underset{\sim}{\lambda}}$ and $\ddot{\underset{\sim}{\lambda}}$. Finally, the method breaks down, through indeterminacy, when the three line-of-sight vectors, the $\underset{\sim}{\lambda}_j$, are linearly dependent, the source of the indeterminacy being the consequent singularity of their matrix. However, we shall find that the angles-only problem, tackled in a different way, is still accurately soluble in the general case of linear dependency and matrix singularity. It is only in the much severer circumstance that the $\underset{\sim}{\lambda}_j$ are dependent as 'localized vectors' that solution break-down is intrinsically inevitable, the implication of this being that the three lines of sight then lie in a single plane which is the orbital plane for an infinite number of solutions to the problem.

2.2 Gauss's method

Gauss's method introduces, directly, a geometric constraint that operates only indirectly (via the dynamics) in Laplace's method: namely, that the vectors $\underset{\sim}{r}_j$ are coplanar, satisfying the relation (say)

$$c_1\underset{\sim}{r}_1 + c_2\underset{\sim}{r}_2 + c_3\underset{\sim}{r}_3 = 0 . \qquad (5)$$

If the c_j were known, independently, we could substitute for the $\underset{\sim}{r}_j$, using (1), to obtain three scalar equations in ρ_1, ρ_2 and ρ_3 (*cf* those, in Laplace's method, in ρ, $\dot{\rho}$ and $\ddot{\rho}$). In reality, of course, the c_j are not known and their eliminant from (5) yields the single scalar equation, of the form $\det(\underset{\sim}{r}_1, \underset{\sim}{r}_2, \underset{\sim}{r}_3) = 0$, that expresses the linear dependence of the $\underset{\sim}{r}_j$. By introducing approximate values for the c_j, however, we can proceed as stated, the basis of the approximation being that c_1, c_2 and c_3 are proportional to $\underset{\sim}{r}_2 \wedge \underset{\sim}{r}_3$, $\underset{\sim}{r}_3 \wedge \underset{\sim}{r}_1$ and $\underset{\sim}{r}_1 \wedge \underset{\sim}{r}_2$ respectively, *ie* to the algebraic values of the triangular areas CP_2P_3, CP_3P_1 and CP_1P_2.

It is, of course, only the ratios of the c's that are needed, and the (approximate) formulae for c_1/c_2 and c_3/c_2 introduce the dynamics, from equation (2) again. Thus, c_3/c_2 can initially be written in the form $-(1 + \gamma_3/r_2^3)t_{12}/t_{13}$, where the term in γ_3 would disappear (by Kepler's second law) if the c's were proportional to swept-out areas (sectors CP_2P_3 etc, as opposed to the triangles). In this way, with the $\underset{\sim}{r}_j$ replaced via equation (1), we get a linear relation between ρ and $1/r^3$ to solve in conjunction with equation (3), just as with Laplace's method. The basic restriction to observations from a single location has entirely disappeared now though, since the dynamic equation results from the elimination of ρ_1 and ρ_3, rather than the artificial $\dot{\rho}$ and $\ddot{\rho}$. In its simplest

form, however, Gauss's method is still limited to short arcs of the orbit, and is certainly not applicable to arcs that cover several revolutions of an elliptic orbit.

To clarify the foregoing, it should be noted that the three equations derived from (5) are effectively in the variables $(c_1/c_2)\rho_1$, ρ_2 and $(c_3/c_2)\rho_3$. After elimination of the first and third variables, and combination with equation (3), solution of the resulting octic (or equivalent) equation leads to values for r_2 and ρ_2. We can now assign values to γ_1/r_2^3 and γ_3/r_2^3, and hence to c_1/c_2 and c_3/c_2; then values (approximate) for ρ_1 and ρ_3, and hence also r_1 and r_3, become available.

The various versions of Gauss's method diverge at this point. The simplest version proceeds via more accurate formulae for c_1/c_2 and c_3/c_2, using the estimates of r_1^3 and r_3^3 (as well as r_2^3) that are now available; the octic equation can be solved again *ab initio*, or the preliminary value obtained for r_2 can be differentially corrected. These formulae (for c_1/c_2 and c_3/c_2) are still only approximations, however, being based on the series modelling of the dynamics relative to the time t_2. To avoid this approximation by series, Gauss devised his celebrated algorithm for the iterative computation of the ratio of the swept-out area to the triangular area for C and any pair of the points P_j. This algorithm, which was summarized by the present author in an appendix of Ref 3, leads to the classical solution of Lambert's problem via the value determined for the appropriate ratio. In the precision version of Gauss's method for the angles-only problem, on the other hand, all three ratios are required for the evaluation of c_1/c_2 and c_3/c_2, so the full solution procedure involves the triple embedding of one iterative process (for Gauss's algorithm) within another (for solution of the overall problem); the orbit is then given by the solution of Lambert's problem with the final values for $\underset{\sim}{r}_1$ and $\underset{\sim}{r}_3$. One advantage from using the precision method is that multirevolution cases can be covered, the revolutions to be included within each arc being carried by the differences in eccentric anomaly. But there remains a serious difficulty with long-arc problems, namely the generation of initial values (for the overall iteration process) that are sufficiently accurate for the process to converge.

Lagrange's and Gibbs's versions of Gauss's method were published, respectively, before and after the version due to Gauss himself. Lagrange's version has features in common with the method of Laplace, his close contemporary, since the orbit is eventually known directly from $\underset{\sim}{r}$ and $\dot{\underset{\sim}{r}}$ (at time t_2), rather than via $\underset{\sim}{r}_1$ and $\underset{\sim}{r}_3$. In Gibbs's version the orbit is modelled by fourth-order approximations that are efficient for hand computation. Also, the parameters of the final orbit are obtained symmetrically from all three of the vectors $\underset{\sim}{r}_j$; the procedure for this essentially elementary problem is outlined in Appendix A.

2.3 Range-iteration methods

The methods of Laplace and Gauss were developed in a very different time from our own: orbits would be restricted to certain types and always heliocentric; only a relatively

short arc of the orbit would be observed; observations would be effectively from a single site; and computing methods would be severely limited. It was a leisurely age, on the other hand, and the computing methods did possess one advantage: the computer (human) could improvise, as necessary, with each problem tackled, and would not be restricted by the constraints of a pre-set 'program'. With the advent of the Space age, however, it was inevitable that methods of greater universality and robustness would be required, whilst the flexibility of operation would have to be built into the software of a machine.

These factors pointed to approaches based on the iteration of estimated values of the ranges, ρ_j, or (equivalently) the distances r_j from the force centre. In the recent paper by Lane[13], for example, the ρ_j are treated as *a priori* independent, so that the iteration is on three unknown variables. This has the advantage of symmetry, but at the expense of complexity; Lane recognized the convergence problem as so acute that he further complicated the problem by the introduction of penalty functions and other sophisticated techniques.

There is no prospect of a method for iterating on a single variable, so consideration can be restricted to two-variable methods*. The methods of Laplace and Gauss may be regarded as of this type, but suffer from the defects that have been noted. We therefore summarize the method described in the text-book of Escobal[10] – it is essentially the same as the method given previously by Briggs and Slowey[12].

We assume that estimates can be made for the distances r_1 and r_3, these being the starting values for the procedure. (Escobal actually works with r_1 and r_2, whilst Briggs and Slowey regard ρ_1 and ρ_2 as the independent unknowns, but the particular preference has no significance.) There is no regular procedure, such as the solution of an octic equation for the methods of Laplace and Gauss, for making the estimates, but 'reasonable' values will be essential if the overall process is to converge to an acceptable solution for r_1 and r_3. Escobal merely suggests that for near-Earth orbits a value 10% greater than the radius of the Earth may be suitable for both variables. (With his three-range method, Lane needs a grid search on the P_j; but for a sufficiently fine grid to capture all solutions, the computer time could be prohibitive, as he notes.)

From the assumed r_1 and r_3, values of ρ_1 and ρ_3 are available via equation (3). (The sign of a square-root has to be selected, a decision that is avoided if ρ_1 and ρ_3 are taken as fundamental, with r_1 and r_3 then the derived quantities.) From ρ_1 and ρ_3, coupled with the observed $\underset{\sim}{\lambda}_1$ and $\underset{\sim}{\lambda}_3$, we now have $\underset{\sim}{\rho}_1$ and $\underset{\sim}{\rho}_3$, and hence $\underset{\sim}{r}_1$ and

* We are assuming that the variables are a pair of ranges, but Dr A.J. Sarnecki has privately suggested that a pair of variables that define the orbital plane, *eg* i and Ω, might be effective; the non-dimensional i and Ω would be unsuitable for Newton-Raphson iteration, however (see section 3.4).

$\underset{\sim}{r}_3$; assuming these to be linearly independent, the location of the orbital plane is then known. If the (localized) vector $\underset{\sim}{\lambda}_2$ does not lie in this plane*, values for ρ_2, $\underset{\sim}{r}_2$ and hence r_2 can now be obtained. (If $\underset{\sim}{r}_1$ and $\underset{\sim}{r}_3$ are linearly dependent, or $\underset{\sim}{\lambda}_2$ lies in the apparent orbital plane, we could try re-starting with different estimates of r_1 and r_3, but it would be more promising to interchange the roles of, say, $j = 2$ and 3 .)

We now have (on the basis of the assumed r_1 and r_3) the polar coordinates of three points in the orbital plane, from which we can derive, by Gibbs's procedure (Appendix A again), the orbital element $p\ (= a(1 - e^2))$ and e , and the true anomalies of the three points. Then the *eccentric* anomalies (elliptic or hyperbolic) are available, and hence the *mean* anomalies, which can be converted to equivalent time differences for the three observations after incorporation of an allowance for integral numbers of revolutions (of an elliptic orbit) when appropriate. But the *true* time differences are known. So we have a pair of residuals, which form the basis for Newton-Raphson iterative convergence for the variables r_1 and r_3. (These residuals are effectively independent, even though the differences t_{12} and t_{23} are not independent when t_{13} is fixed, because Gibbs's procedure does not use the time differences.)

Once the correct values of the r's are available, with the final values of p , e and eccentric anomalies as a by-product, it is a routine matter to complete the solution of the problem by the computation of $\dot{\underset{\sim}{r}}_2$.

The Newton-Raphson convergence process requires the partial-derivative matrix of the dependent variables (pair of time differences) with respect to the two independent variables. Escobal specifies the computation of these derivatives numerically, by recalculation of the time differences after incrementing first r_1 and then r_3 . For the increments he suggests 4% of r_1 and r_3 themselves, values which seem (to the present author) far too big: since the derivatives are 'one-sided' (see also section 3.3), the resulting effect on truncation error is potentially very serious. Escobal also suggests, for near-Earth orbits, a value for the tolerance to be used as the criterion for the completion of convergence. This topic, as with the increments for partial derivatives, will be considered in the context of the new method to be presented in the next section.

An important matter that is ignored in both Refs 10 and 12 is the handling of multiple solutions, the danger with an unlucky choice of the pair of starting values being that convergence will be to a 'wrong' solution, rather than that it will fail altogether. Lane[13],

---

* This restriction on $\underset{\sim}{\lambda}_2$ arises from the approach of Refs 10 and 12, and is not inherent in two-range iteration. The trouble arises from the purely geometric derivation of ρ_2 and r ; it is avoided (as we see in the next section) by bringing in the dynamics via Lambert's problem. There is no difficulty anyway if just O_2 lies in the orbital plane, the implication then being that $\rho_2 = 0$.

however, emphasizes the need to compute all solutions, if possible, and this is a philosophy that the present author strongly supports (see section 3.5).

3 THE NEW METHOD

3.1 Preliminary remarks

The new method is a range-iteration one, but is only broadly in the same family as the method just outlined. The fundamental difference is that, whereas Escobal's method employs the two time differences (relative to t_2) as 'target functions', a procedure based on the new method computes the body's position at t_2 from the orbit derived on the basis of the assumed positions at t_1 and t_3 ; then the target functions are the projections of this computed position on a plane perpendicular to the known sight-line at t_2 . The rationale for this virtual inversion of Escobal's procedure is that it permits the utilization of the author's robust and accurate procedure for Lambert's problem[3], which always converges and gives at least 13-digit accuracy after three iterations of a computation based on a single parameter. (See section 5 for further comparison of the two approaches.) Particular attention had been paid to awkward cases in the Lambert procedure, including the allowance for multiple revolutions and the maintenance of full accuracy and rapid convergence in situations of marked nonlinearity; it was believed that the Lambert approach to nonlinearity would minimize any difficulties from this source in the angles-only procedure. Another ready-made tool for the new method was provided by the author's technique[14] for propagating state vectors (position and velocity) via the set of 'universal' elements $\alpha, q, i, \Omega, \omega$ and τ, where $\alpha = \mu/a, q = a(1-e)$ and τ is the time at a pericentre; the minimal-error Kepler solutions[2] constitute an important component of this technique.

The next subsection describes the basic method, and the following subsections give further details of some important aspects, the complication associated with multiple revolutions being the last to be dealt with. Then section 4 is devoted to the behaviour of the implementing procedure (on a particular computer) in practice.

3.2 Basic method

It is recalled that $\underset{\sim}{R}_j$ and $\underset{\sim}{\lambda}_j$ at t_j (for $j = 1, 2$ and 3), together with μ, constitute the data, but we shall also require an assumed value for k, the number of half-revolutions included in the angle P_1CP_3 – only for elliptic orbits can k exceed 1. (The right value for k will sometimes be known; otherwise, we must run the procedure with several values if necessary.) It has been assumed that the $\underset{\sim}{\lambda}_j$ are unit vectors, such that $\underset{\sim}{\rho}_j = \rho_j \underset{\sim}{\lambda}_j$, but in practice they may not (as given) be exactly (or even approximately) 'unit', in which case the unique oriented unit vector, $\hat{\underset{\sim}{\rho}}_j$, could be distinguished from the given positive multiple of it, $\underset{\sim}{\lambda}_j$; to avoid this complication (taken account of in the actual computing procedure), however, we continue to identify $\underset{\sim}{\lambda}_j$ and $\hat{\underset{\sim}{\rho}}_j$. It should be noted that these (unit) vectors

refer to lines having a definite location in space (defined by the points O_j, *ie* by the $\underset{\sim}{R}_j$) as well as direction, and this 'affine' property (as opposed to a purely vectorial one) will routinely be implied in references to the $\underset{\sim}{\lambda}_j$, making them 'localized' vectors rather than just 'free' vectors (*cf* the remark at the end of section 2.1). The *specific* positions of the O_j on the lines are almost irrelevant, however; this will be referred to as 'the principle of observer independence' and it is desirable that a solution procedure reflects this principle as far as possible*.

The foregoing remarks are relevant to the definition of our 'target functions', f and g, which is based on the sight-line associated with $\underset{\sim}{\lambda}$ (= $\underset{\sim}{\lambda}_2$). This line, being oriented, may be identified with the normal to a family of parallel oriented planes, and its intersection with each plane may be taken as the origin of a coordinate system within that plane. In each iteration of the solution procedure we compute a position, P_c, at time t (= t_2), and this identifies (unambiguously) the plane of the aforesaid family that passes through P_c. Our target functions are simply the coordinates (f,g) within this plane; since the origin (P_c) has been specified, it remains to be specific about the directions of the axes of f and g. The following explanation should be clear if it is borne in mind that the axis system (like the plane in which it is embedded) is defined afresh for each iteration of the convergence process.

So long as P_c does not lie precisely on the line $\underset{\sim}{\lambda}$, we can orient our axes (within the oriented plane) uniquely, and almost magically, just by specifying $g = 0$ and $f > 0$. The 'magic' may be dispelled by noting that, though the axis system is ephemeral, it can temporarily be thought of as permanent! Thus for different values of ρ_1 and ρ_3, which we will re-denote by x and y for convenience in the function analysis, different values of both f and g would result, so that well-defined and non-trivial values exist for partial derivatives such as (we denote by) f_x and g_{yy}. This orientation of axes breaks down, of course, if P_c lies exactly on $\underset{\sim}{\lambda}$ (so that $f = g = 0$), but we can only have such a P_c when a precise solution of the angles-only problem has already been obtained, so that we do not have to proceed any further; the important point here is that there is no loss of computing accuracy as this 'singularity' is approached.

A considerable advantage over other range-iteration methods is worth repeating (from an earlier footnote) at this point: whereas Escobal's method breaks down if the vector $\underset{\sim}{\lambda}_2$ lies in the orbital plane, this is of no consequence whatever with the new method. Linear dependence of the vectors $\underset{\sim}{r}_1$ and $\underset{\sim}{r}_3$ at any stage will still cause failure, but this

---

* The principle is not reflected at all in the methods of Laplace and Gauss and only partly in Escobal's; the actual O_j are only relevant because the orientations attached to the $\underset{\sim}{\lambda}_j$ are known in practice (see the distinction, in an earlier footnote, between 'directions' and 'anti-directions').

can be dealt with by an interchange of data, say $j = 2$ with $j = 3$, as suggested in section 2.3, unless of course $\underset{\sim}{r}_2$ lies in the same direction as both $\underset{\sim}{r}_1$ and $\underset{\sim}{r}_3$, in which case there is a *fundamental* indeterminacy. (It may be that one solution, for the given data, is indeterminate in this way, though other solutions exist that are entirely well-behaved.)

The overall structure of the new procedure, written in Fortran-77, is that the main program calls OBS3LS, the subroutine responsible for generating each individual solution of a given angles-only problem, whilst OBS3LS calls CALCPS, a subroutine implementing a self-contained algorithm which uses estimated values for x ($= \rho_1$) and y ($= \rho_3$) to calculate $\underset{\sim}{c}$, a version of the vector $\underset{\sim}{\rho}_2$ that is independent of the actual observation $\underset{\sim}{\lambda}_2$. The detailed operation of CALCPS is described in Appendix B, whilst Appendix C provides a listing of both CALCPS and OBS3LS. An outline of the operation of OBS3LS is given in the next paragraph. The main program has the following functions (among others): to register the success or failure of a given solution attempt (by OBS3LS) for a given value of k; after a success, to search for a further solution if required (see section 3.5); and to organize two separate sets of solutions if $k \geq 2$ (see section 3.6).

The subroutine OBS3LS has four dummy arguments, the first two of which are purely input arguments, viz NHREV (the value of k) and IND (an indicator that distinguishes between the two sets of solutions just referred to). The other two arguments specify ρ_1 and ρ_3, initial estimates (starting values) of which must be supplied as input; these estimates are updated iteratively by the subroutine, and successful operation leads to a pair of corrected values that constitute a solution to the problem. The current values of x and y are used in each iteration, as inputs to CALCPS, to provide $\underset{\sim}{c}$, from which OBS3LS derives f (with $g = 0$). It continues (within each iteration) with the evaluation of the relevant partial derivatives of f and g via five additional CALCPS calls with incremented values of x and y as appropriate (see section 3.3 for details of this evaluation and for the formulae that use these derivatives as the basis for the update of x and y during each iteration). After the new values for x and y have been obtained in the current iteration, a test for 'convergence complete' is made (see section 3.4 for consideration of starting estimates and of convergence completion); if convergence is not complete, a further iteration (to a limit of 25, which is arbitrary but easily altered) takes place, beginning with the use of CALCPS to derive a new $\underset{\sim}{c}$ vector, and hence (with new axes for f and g) a new value for f. The convergence test is based on the value of f obtained *before* the refinement of x and y, so the refinement during the final iteration is often superfluous. Consideration was given to making the procedure more efficient by placing the test at the *beginning* of each iteration, rather than the *end*, but there were a number of reasons for not doing this: in particular, the convergence criterion could be relaxed (so that in no case would it be impossibly severe) in the confidence that best-possible values of x and y would nevertheless be produced; this often avoids a superfluous iteration.

It has been noted that OBS3LS requires a separate call for each solution of a given angles-only problem. The necessary information regarding solutions already derived, if any, is held in the common block /KNWNSL/, and further details of the operation of the subroutine from this point of view are given in section 3.5. (The first two arguments will be the same for all the solutions of a given set, but the other two arguments, ρ_1 and ρ_3, must be reset from the solution values just obtained to the starting estimates required for the next solution.)

When $k \geq 2$, two sets of solutions are sought. The problems associated with multi-revolution coverage are considered in section 3.6.

3.3 Iteration formulae and partial derivatives

The classical Newton-Raphson formula for iterative refinement of the root of an equation generalizes easily to any number of variables, assuming the same number of equations. The iteration formula for two variables, in particular, is

$$\begin{pmatrix} \delta x \\ \delta y \end{pmatrix} = - \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}^{-1} \begin{pmatrix} f \\ g \end{pmatrix}, \tag{6}$$

where δx and δy are the corrections to the current estimates of one pair of roots of the equations $f(x,y) = g(x,y) = 0$. Here we assume $g = 0$ already, so equation (6) reduces to

$$\delta x = - D^{-1} f g_y \tag{7a}$$

and

$$\delta y = D^{-1} f g_x , \tag{7b}$$

where

$$D = f_x g_y - f_y g_x , \tag{8}$$

ie D is the Jacobian (determinant of the derivative matrix).

These simple iteration formulae, involving the partial derivatives of f and g with respect to x and y, are obtained by truncating the Taylor expansions of f and g after the linear terms. Then (just as with a single variable) iteration converges to a solution whenever starting values (for x and y) are available that are not too remote from that solution. So long as the two terms of D are not nearly equal, moreover (*ie* the derivative matrix is well conditioned), the convergence is quadratic, reflecting the increasing legitimacy of the truncation as the solution is approached. Ideally we would obtain the partial derivatives (f_x, g_x, f_y and g_y) analytically, in particular by extending the CALCPS algorithm to provide derivatives for all quantities. There has been no serious attempt to obtain analytical

derivative formulae, however, as experience has indicated that numerical derivatives are accurate enough for the angles-only problem.

The most economical derivation of the four partial derivatives, numerically, requires only two additional CALCPS calls in each iteration: in the first, x is given a suitable increment, so that f_x and g_x can be obtained (it is recalled that the axis-system is not re-defined within an iteration, so a non-zero value of g will result, in general, from this increment); in the second additional CALCPS call, it is y that is incremented, so that f_y and g_y can be obtained. But the derivation of f_x, say, from $[f(x + \delta x) - f(x)]/\delta x$, where the argument y has been suppressed, involves an unfortunate bias, arising from the one-sided increment, and derivation via the formula $[f(x + \delta x) - f(x - \delta x)]/2\delta x$ would clearly be much more satisfactory. To get our derivatives this way, we have to increase the number of additional calls from two to four, but as a bonus we then immediately have estimates for the second-order derivatives f_{xx}, g_{xx}, f_{yy} and g_{yy}, where the first of these is given by $[f(x + \delta x) + f(x - \delta x) - 2f(x)]/(\delta x)^2$. Thus to get a complete set of second-order derivatives, we only require a total of five 'additional' calls to CALCPS, the fifth (and final) being required for f_{xy} and g_{xy}. This final pair of derivatives will necessarily be biased, but the use we make of f_{xy} and g_{xy} is sufficiently marginal for this to be unimportant.

Second derivatives are useful because they permit the Newton-Raphson iteration process, with quadratic convergence, to be superseded by a process that gives cubic convergence. As a generalization of the corresponding process for a single variable, we may refer to this as Halley's process. (Ref 1 provides an account of a number of root-finding processes for a single variable, of which the basic Halley method is one, the application being to Kepler's equation.) To derive the formula for Halley's process, we note that (with third-order derivatives neglected) we want δx and δy to satisfy

$$f + f_x\,\delta x + f_y\,\delta y + \tfrac{1}{2}f_{xx}(\delta x)^2 + f_{xy}\,\delta x\,\delta y + \tfrac{1}{2}f_{yy}(\delta y)^2 = 0 \qquad (9)$$

and the corresponding equation for g. But we have the Newton-Raphson values for δx and δy, as given by equations (7), and these can be used to linearize (9) and thereby obviate the need to solve a pair of simultaneous quadratic equations. Thus we replace $(\delta x)^2$ and $(\delta y)^2$ by $(\delta x)_{NR}\,\delta x$ and $(\delta y)_{NR}(\delta y)$, respectively, just as with a single variable[1]; there is an arbitrary aspect to the corresponding replacement of $\delta x\,\delta y$, but replacement by $\tfrac{1}{2}[(\delta y)_{NR}\,\delta x + (\delta x)_{NR}\,\delta y]$ is the obvious choice, based on the replacement of

$$\tfrac{1}{2}(\delta x \quad \delta y)\begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}\begin{pmatrix} \delta x \\ \delta y \end{pmatrix} \quad \text{by} \quad \tfrac{1}{2}(\delta x_{NR} \quad \delta y_{NR})\begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}\begin{pmatrix} \delta x \\ \delta y \end{pmatrix}.$$

Thus we replace (9) by an equation that (because the third derivatives were already being neglected) is no less valid, viz

$$(f_x + \tfrac{1}{2}f_{xx}\,\delta x_{NR} + \tfrac{1}{2}f_{xy}\,\delta y_{NR})\,\delta x + (f_y + \tfrac{1}{2}f_{xy}\,\delta x_{NR} + \tfrac{1}{2}f_{yy}\,\delta y_{NR})\,\delta y = -f, \quad (10)$$

with the corresponding equation for g. Evidently, the Halley iteration formula is just (6) with the partial derivatives modified in line with the coefficients in (10).

However, another useful possibility exists, just as with a single variable[1], viz to replace each occurrence of "½" in (10) by "1". The resulting formula works better than Halley's formula when two or more solutions for (x,y) are close together relative to the starting point. (The one-variable formula is obtained on applying the basic Newton-Raphson formula to f/f' rather than f, *ie* to a function that has only a single root where f has a repeated root; the formula based on (10) can be shown to be the natural extension to two variables.) Thus a general procedure might be to start iterating with the 'modified Newton-Raphson' process, *ie* with the ½'s replaced, and then switch to the Halley process. Subroutine OBS3LS has not been written to work like this, but it *has* been arranged (via the Fortran variable HN held in common block /OTHER/) that both modes of operation are possible (or other modes if HN is given a value other than 0.5 or 1.0); further, at most 20 iterations are allowed under 'modified Newton-Raphson', after which the Halley process automatically takes over. For a test example with comparison of the two processes, see section 4.2.

Finally, a rare circumstance has been identified (see section 4.5(b)) in which progress can be made with the geometric mean of the pure Newton-Raphson δx and the Halley δx (and likewise for δy), though neither is helpful on its own. The situation is essentially the opposite of the preceding one (in which two solutions are close together relative to the starting values), arising instead when the starting values are midway between two solutions. As can be seen very clearly for a single-variable problem, such as the iterative solution of the equation $x^2 = 1$ from a starting estimate close to zero, what happens is that the Newton-Raphson increments are extremely large, as a result of which, by equation (10), the Halley increments are correspondingly small and hence useless. But very large Newton-Raphson estimates are caused by a (numerically) very small value of D, which is itself associated with near singularity of the derivative matrix. Thus the situation can be recognized from the value of $|D| / (|f_x g_y| + |f_y g_x|)$ and a late amendment was made to OBS3LS such that if this quantity (given by the Fortran variable DELNM) is less than 0.001, then the Halley increments are replaced by the pair of geometric means (ignoring here the complication that arises if either is imaginary). It can be inferred from the pair of quadratic equations why it is that the geometric means are of the right order of magnitude and work better than either pair of normal increments in the given circumstance; even better (of course) would be an exact solution of these simultaneous equations.

A serious risk exists with any root-finding procedure (or with any minimization procedure[15] based on the use of partial derivatives) that, though a 'step' is taken in a desirable direction, the step is too long and does more harm than good. With the angles-only problem, via the approach being described, it was sometimes found that $f\,(>0)$ increased after an iteration, instead of getting smaller. The supposition that in most cases this involved an 'overshoot' was confirmed when it was found that the new value of f would have had a negative sign attached if it had not been for the re-orientation of axes. As an empirical treatment of the phenomenon, the following action was incorporated in the procedure whenever the value of f is more than doubled from one iteration to the next: if the superseded values of x and f are denoted by $x_{\rm old}$ and $f_{\rm old}$ ($< \frac{1}{2}f$), then the new value of x (and similarly y) is replaced by $(f\,x_{\rm old} + f_{\rm old}\,x)/(f + f_{\rm old})$; it will be seen that this is in conformity with the secant or *regula falsi* method for one variable if the assumption above, concerning the effect of axis re-orientation on the sign of f, is correct. This additional action takes place at the beginning of the new iteration and is immediately repeated if the revised f still exceeds $2f_{\rm old}$. After that, the new iteration is resumed regardless of $f/f_{\rm old}$.

In regard to the numerical evaluation of the partial derivatives, it remains to consider the values adopted for the increments, $\pm\delta x$ and $\pm\delta y$, to x and y in the 'additional calls' of the subroutine CALCPS. The problem here is that δx and δy must not be too small, or unacceptable rounding error in subtractions such as $f(x + \delta x) - f(x - \delta x)$ will occur; equally they must not be too large, or the truncation error will be too great. The first (arbitrary) choice was a simple 1% one, with $\delta x = 0.01x$ and $\delta y = 0.01y$, using the current values of x and y (where it will be recalled that $x = \rho_1$ and $y = \rho_3$). For a number of reasons (see the end of section 3.6 for one of these) it was soon decided that the choice of 0.1% would be better. However, neither choice took account of the desire, referred to in section 3.2, to make the procedure as 'observer-independent' as possible, *ie* with results independent of the positions of the points O_j on the lines $\underset{\sim}{\lambda}_j$; in particular, the derivation of the trivial solution for heliocentric orbits (see section 2) would be prone to severe error for $x \approx y \approx 0$. The resolution of this difficulty was straightforward, though at the expense of some efficiency: instead of taking $\delta x = 0.001\ \rho_1$, we actually take $\delta x = 0.001\ r_1$ (and $\delta y = 0.001\ r_3$), the values of r_1 and r_3 being observer-independent.

It will be noted that the increments used as $\delta\rho_1$ and $\delta\rho_3$ will become effectively constant as the solution is approached, and it might be thought that there would be an advantage in having them diminish as the iterative corrections themselves diminish. This possibility was investigated, with the increments used in each iteration (after the first) limited by the corrections made in the previous iteration. Little effect was found on convergence, but the tendency seemed to be marginally harmful rather than beneficial. The good behaviour of unrefined derivatives may be attributed to the fact that the computation of

first derivatives using two-sided increments eliminates truncation error due to second derivatives.

3.4 Starting values and convergence criterion

As with so much of the angles-only procedure, there is inevitably a considerable degree of arbitrariness associated with both the initiation and the termination of the iteration process. The former is expressed by the starting estimates for the values of ρ_1 and ρ_3 (we now abandon the notation x and y for these quantities), and the latter by the particular test selected as the criterion for the completion of convergence.

In regard to starting values, it was considered desirable to set these outside the subroutine (OBS3LS) for iterating the solution, so that various options could be covered; in particular, they can be set on the basis of assumed knowledge of the type of orbit, or else given default values; by whatever method, this is virtually the only point of the overall procedure at which we have to relax the observer-independence principle. It is enough here just to state the default values that have been used in the procedure: on a basis that is almost totally arbitrary, both ρ_1 and ρ_3 are set to $2(R_1 + R_2 + R_3)$, which is certainly observer-dependent. Additional options arise when multiple solutions are sought (see section 3.5).

The convergence test involves two things, the quantity actually tested and the numerical criterion value. An obvious quantity to test was the non-dimensional* f/c, where it will be recalled that $\underline{c}$ is the calculated version of $\underline{\rho}_2$, but this is not always satisfactory and a more carefully chosen quantity is needed. First, however, we consider the criterion value, and here it was decided that (until the solution procedure could be regarded as an operational tool) accuracy was more important than efficiency. In consequence the extremely conservative criterion of 10^{-12} was adopted, related to the 16-digit double-precision accuracy (IEEE standard) available with the Amstrad PC1640 computer on which the procedure was developed. (If the test were conducted at the beginning of the iteration, this would not be conservative at all; but when a process is providing cubic convergence, and the test is at the end, it implies that, but for the inevitable rounding error, a relative precision of 10^{-36} could be attributed to the results of the final iteration!)

Returning to the quantity tested, which will be denoted by ε, we first note that the problem with f/c is not that it is observer-dependent but that it leads to a ludicrously stringent test when c is (in relative terms) very small, as with the trivial solution for

---

* Though the *testing* of a non-dimensional quantity is appropriate, the use of non-dimensional *target functions* (f/c and g/c) would be disastrous in a Newton-type convergence process – this point is explained in section 3.5.

heliocentric orbits again. We avoid this difficulty, very easily, on replacing the denominator, c, by the larger of c and (the calculated version of) R; R is more convenient than the more obvious r, their effects differing by at most a factor of two, after the MAX operation. A more subtle point arises with highly elliptic orbits, however, since the modified test may still be unrealistic when the body has been moving rapidly near its pericentre (see the discussion on testing, in similar circumstances, in Refs 3 and 14). This has been dealt with, semi-empirically, by introducing $V_2 t_{12}$ into the MAX operation, where V is velocity.

3.5 Multiple solutions

We know that, even for a fixed value of k (in particular for $k = 0$ or 1), it is common for there to be more than one solution to the angles-only problem. Thus, we have seen (in section 2) that the general number of solutions for short-arc coverage is three. It is assumed in this section that our purpose is to obtain as many solutions as possible for each value of k that is considered relevant, so there is no question of resting on our laurels after a first solution has been obtained. To obtain a second solution (and then, maybe, a third) there are two obvious approaches: to alter the starting values and hope that this is enough for the same iteration process to produce different results; or to alter the iteration process itself, in such a way as to steer away from the known solution. We will take the second approach, and for simplicity proceed on the assumption that the starting values are unchanged. There is no reason why different starting values should not be used, however, and a particular option for revised values has been built into the overall procedure that calls the subroutine OBS3LS listed in Appendix C; this option is described at the end of the section.

To 'steer away' from one or more 'known solutions' we replace the target functions, f and g, by related functions, F and G say, which have the same *unknown* roots as f and g do. At the known root, or roots, we would like F to be non-zero in principle, perhaps even infinite or indeterminate; a zero-value would be permissible, but only if the attractive power of this zero is sufficiently diminished. (G is necessarily zero, or indeterminate, because of the way in which G is defined.) To get from f and g to F and G, therefore, an obvious possibility is to divide by a quantity that vanishes at the known root(s), but we shall find that this is not sufficiently subtle on its own.

Returning to the notation (x,y) for (ρ_1,ρ_3), we introduce (X,Y) to designate a known solution, and write

$$(\xi,\eta) = (x - X, y - Y) . \tag{11}$$

Then if

$$\beta = \sqrt{\xi^2 + \eta^2} , \tag{12}$$

we could work with f/β as our desired F, and g/β as G. The use of the square root, which gives β the dimension of a distance, is appropriate for 'flattening' f out, but the implicit value of 1.0 as an exponent, when β is actually used, could be replaced by some other value. (It is easy to make this change in OBS3LS, via a change to the quantity ONE, set to 1D0, and other values have been tried; the effect on convergence was bad, however, perhaps not surprisingly, even with quite a small variation from 1.0.) If we already have two known solutions, then we would set F to $f/\beta_1\beta_2$, say, so there is no additional difficulty in principle. Before proceeding, we note that β is observer-independent, since ξ and η are unchanged if x and X (and similarly y and Y) change by the same constant.

The problem with f/β as F is that it generates a function of entirely the wrong form for convergence by a Newton-type process. The form we want must not be too different from a polynomial (in two variables), and in particular we want $f \to \infty$ as $x \to \infty$ or $y \to \infty$. Put another way, the objection to f/β is that the dimensional character of f has been cancelled out. So to 'cancel this cancellation' we actually set

$$F = f\gamma/\beta, \tag{13}$$

where γ has the same dimension (length) as β and f; moreover, we choose γ so that it is always positive and (like β) is observer-independent. This still leaves a considerable degree of arbitrariness, and in practice the choice

$$\gamma = \sqrt{\beta^2 + r_{1,\mathrm{kn}}^2 + r_{3,\mathrm{kn}}^2} \tag{14}$$

was made, where $r_{1,\mathrm{kn}}$ and $r_{3,\mathrm{kn}}$ are the (fixed) values of r, for the known solution, at times t_1 and t_3.

We require the values of the partial derivatives F_x etc, but these can be obtained at once from the (numerical) values, f_x etc, that have been derived by the basic procedure. All we need to get from the f-derivatives to the F-derivatives are the first derivatives of β and γ, which are given by

$$\beta\beta_x = \gamma\gamma_x = \xi \qquad \text{and} \qquad \beta\beta_y = \gamma\gamma_y = \eta\ . \tag{15}$$

It then follows at once that

$$F_x = (\gamma/\beta)(f_x - w\,\xi f)\ , \tag{16}$$

where

$$w = 1/\beta^2 - 1/\gamma^2\ , \tag{17}$$

and similarly for F_y, G_x and G_y (with $g = 0$ as usual). Before considering F_{xx} etc, it is remarked that we do not have to generate F_x etc explicitly, since it follows from the homogeneous form of equations (7), after substitution for D, that it is more convenient just to modify f_x etc, writing

$$f_{x,\mathrm{mod}} = f_x - w\,\xi f \quad . \tag{18}$$

For the second derivatives it now follows that we merely have to modify f_{xx} etc in conformity with

$$f_{xx,\mathrm{mod}} = f_{xx} - 2w\,\xi f_{x,\mathrm{mod}} - wf[1 - \xi^2(1/\beta^2 + 3/\gamma^2)] \tag{19}$$

and

$$f_{xy,\mathrm{mod}} = f_{xy} - w\,(\eta f_{x,\mathrm{mod}} + \xi f_{y,\mathrm{mod}}) + w\,\xi\,\eta f\,(1/\beta^2 + 3/\gamma^2)\ . \tag{20}$$

It is fairly straightforward, as already intimated, to extend the modification of the derivatives when the number of 'known solutions' allowed for is greater than one. It is also possible to modify all the formulae to allow for a non-unit exponent to be attached to β or γ, but (again as already remarked) this has been found unrewarding.

Since we are merely aiming to steer the convergence away from existing solutions, it would be natural to return to unmodified derivatives after (say) two or three iterations with modified derivatives, but this complication has not been introduced in the present version of the procedure. To make the convergence criterion the same for all solutions, however, the test is always on the unmodified f, not F, in spite of the risk (considered negligible) that a known solution might then be rediscovered. It would not be necessary to evaluate F at all, in fact, except that the overshoot-avoidance device, if used, must be in terms of F.

We now define the option (referred to earlier) for revising the starting values for solutions after the first – it has been found more successful, in practice, than re-use of the original starting values. The starting values for the second solution are defined by $(2X_1 - X_0, 2Y_1 - Y_0)$, where (X_0,Y_0) and (X_1,Y_1) denote the original starting values and the first solution respectively. The starting values for the third solution are defined, as $(2X_2 - X_1, 2Y_2 - Y_1)$, by the first two solutions, and this use of the last two solutions continues if yet further solutions are sought.

In subroutine OBS3LS, information about the known solutions is taken from the common block /KNWNSL/; the number (up to nine) is provided by the integer variable IKN, which must be set to the appropriate value (zero initially) before each solution to a given problem is sought. The known ρ_1 and ρ_3 themselves are assumed to be in the arrays RO1KN and RO3KN, whilst R1KNSQ and R3KNSQ hold $r^2_{1,\mathrm{kn}}$ and $r^2_{3,\mathrm{kn}}$.

3.6 Multirevolution coverage

We now cover situations where the interval between t_1 and t_3 covers more than one complete orbital revolution, so that the orbit is necessarily elliptic. It is possible, in fact, that a solution (perhaps more than one) exists for each value of k up to some maximum value. The following argument shows that the limiting value is finite: the greater the number of revolutions that we seek to squeeze in between t_1 and t_3 (as k increases), the smaller must be a, the semi-major axis; but there is a non-zero lower bound to a, since otherwise the three lines of sight, the $\underset{\sim}{\lambda}_j$, would all pass through the force centre, and this condition ('all observations at the zenith' in the case of observations from a spherical planet) can be ruled out on the basis that it implies either inconsistency or indeterminacy. (See also the remarks in Appendix A and the footnote within section 4.2.)

It has been noted that the solution of Lambert's problem over the interval (t_1,t_3) lies at the heart of the CALCPS algorithm, and hence at the centre of the angles-only procedure itself. But there is a crucial difference between Lambert solutions for $k \leq 1$ and $k \geq 2$, as follows from Ref 3 and the plot 'of T vs x' within that paper. Thus, for $k = 0$ or 1 there is always exactly one solution to a given Lambert problem. For $k \geq 2$ (the multirevolution cases), on the other hand, there will either be two solutions or else, when the implicit energy is too great, no solution; Ref 3 also covers the limiting case of a single solution, where two coalesce, but we can disregard this as being of zero probability in practice. (Computational difficulties associated with two close Lambert solutions cannot be so easily dismissed, but this issue, more complicated than the related one for $k \leq 1$, has not been addressed.)

For a given value of k (≥ 2), and for each solution sought, OBS3LS proceeds (as noted in section 3.2) via calls to subroutine CALCPS (see Appendix B); these calls indicate, via the argument IND, which of the two Lambert solutions (if they exist) is to be used. Since IND is also an argument (otherwise unused) of OBS3LS itself, to find all solutions of the angles-only problem two sets of calls to this subroutine are required, with IND set to unity for each solution in the first set and then to zero for each in the second set (zero designates the Lambert solutions of lower energy and unity those of higher energy). Thus when $k \geq 2$ there are considered to be two entirely separate cases, each of which (as uniquely for $k \leq 1$) can lead to a set of several solutions or none; the second case is looked at independently of failure with the first case, an eventual such failure being inevitable, of course, if too many solutions are looked for. An unfortunate new source of failure arises with the search for solutions of either set, however, as a direct corollary of the 'no solutions' possibility for the Lambert problem at any stage. Thus a solution may be under way with a value of k that is actually the 'correct' one, when the value of ρ_1 and ρ_3, as generated by a particular iteration, are such that the next iteration breaks down; the breakdown can occur in the first iteration if the energy implied by the starting values (and the value of k) is too high.

A fully satisfactory way to deal with the general situation has not been found, but some fair success has been obtained with the following modes of operation, which have been built into OBS3LS. If there is an immediate failure with the given starting values, then these are replaced by the values of ρ_1 and ρ_3 corresponding to the common perpendicular (shortest transversal) to the two lines of sight; if either of these is negative, it is replaced by zero. If there is still an immediate failure, then a final attempt is made to get the iteration going by setting both ρ_1 and ρ_3 to zero. If a failure occurs at the beginning of the second (or a subsequent) iteration, on the other hand, then it is postulated that ρ_1 and ρ_3 were over-corrected and the increments just made are replaced by increments of one-third their value; if this does not work, two further such reductions in the increments are tried. After three failures, the process is abandoned as if the failure occurred on the first iteration, whereupon there is recourse to the common-perpendicular or zero starting values. Failures are registered by means of the Fortran variable NFAIL.

A point of minor irritation was met on a number of occasions, namely, that a Lambert failure would be encountered in one of the CALCPS calls associated with incrementation for partial derivatives (even after the change from 1% to 0.1%, made partly for this reason), though not in the main call. This was dealt with by a reduction of the increment, for one iteration only, by a factor of 10; should that not suffice, up to two further reductions are tried (with a final overall reduction by 1000) before the process is abandoned. The effect of this modification was the virtual elimination of this problem.

4 TEST RUNS

4.1 An example of Herrick

The textbook of Herrick[11] is probably unique (in the field of orbital dynamics) for the thoroughness of its presentation. In the numerical examples, in particular, he presents the result of every stage of the calculation, which makes these examples highly suitable for comparison purposes. To illustrate Laplace's method (Chapter 12 of Ref 11) and Gauss's method (Chapter 13), he works with three observations of the minor planet 683 Lanzia (1909 HC), having previously used data for this body to exemplify his algorithms for ephemerides based on position and velocity (Chapter 7). We take, as data for the $\underset{\sim}{R}_j$ and $\underset{\sim}{\lambda}_j$, the numbers he lists (with distances in *astronomical units*) as part of 'Example 12C1', together with the times t_{12} and t_{13} in conventional units to permit adoption of the value $\mu = 1$; these quantities are provided here as Table 1. (All the numbers, curiously, are positive, though Herrick gives all the $\underset{\sim}{R}_j$ components as negative since his definition involves the vectors O_jC rather than CO_j as here.)

We will not quote Herrick's own solution for ρ_1, ρ_2 and ρ_3 , for a number of reasons: his results are (unsurprisingly) not quite the same for Gauss's method as for

Laplace's; we will not be incorporating light-time corrections, so that Herrick's results would not be what we expect to get in any case (the adjusted values of t_{12} and t_{13}, obtained in Herrick's solution by Gauss's method, are included here in Table 1); the new procedure operates to many more figures than Herrick worked to; and the new results are self-checking on the basis of the final value of the convergence criterion for each solution obtained. It may, however, be of interest to list all the real roots of the octic equation that Herrick effectively sets up (as part of 'Example 13C3') in Gauss's method – the octic in Laplace's method has significantly different coefficients. The four roots for r ($= r_2$) are 3.239..., 0.963..., 0.875... and -3.368... . Only the positive roots are relevant, and these correspond to values for ρ of (respectively) 2.563..., -0.040... and -0.926 . Herrick derives only the first root, which leads to the correct solution of the problem, the others being not merely wrong (for the actual minor planet) but invalid. Here we will give all three solutions obtained by the new method, only the first being consistent with Herrick's octic equation; the final solution is an example of the trivial solution for heliocentric orbits observed from the Earth.

The three solutions with $k = 0$ are the only ones. They are listed in Table 2, each solution being defined by ρ_1 and ρ_3, with ρ_2 given for completeness. The default starting values (5.9... for both ρ_1 and ρ_3) were used for the first solution, the other two being generated, with revised starting values, as described in section 3.5. The number of iterations to generate each solution is tabulated as N, whilst ε gives the final-iteration value of the test quantity that must be less than 10^{-12} for convergence to be deemed to have occurred; finally, e is the approximate eccentricity of the orbit.

The maximally significant number of figures* is given for each ρ-entry, following their derivation on the computer used, viz an Amstrad working to 16 decimal digits as already noted. The values of ρ_2 are intrinsically less accurate than for ρ_1 and ρ_3, since each iteration terminates with the latest ρ_1 and ρ_3 for which the corresponding ρ_2 is only derived if there is a further iteration – the full accuracy for ρ_2 was easily obtained, however, without recompilation of the program to force an additional iteration, by running with the built-in option for interchange of the second and third observations.

The first solution has evidently been accompanied by an accuracy loss of no more than three decimal digits due to rounding error, and the accuracy of the second solution is even better, which is perhaps surprising; the third solution is as good as the second in absolute terms, the loss of three more significant figures in *relative* terms being the inevitable consequence of the small values involved. The accuracy of the solutions, and the

* The precision of an angles-only solution is easily assessed via a sample of runs with different starting values, or by running with the data-interchange option.

rapidity of convergence, can largely be attributed to the excellent condition of the derivative matrix in the vicinity of each solution.

If the same (default) starting values are used for all three solutions, the second and third emerge in reverse order, the number of iterations for the three being 4, 11 and 17. If the progress of each solution is plotted in the plane of $(x,y) = (\rho_1,\rho_3)$, it appears that the reduced speed of convergence to the second and third solutions may be attributed to the fact that the iteration process must effectively pass through the first solution to locate what has now become the second, and then through both known solutions to locate the final one – with 'revised' starting values this phenomenon only occurred with the third solution, and in a milder form. It is remarkable, on the other hand, how robust the entire process turned out to be when experiments with different numbers for the default starting values were made, whether with revision (for the second and third solutions) or not. Even with enormous values, *eg* with $\rho_1 = \rho_3 = 1000000$ (which takes us more than 25000 times beyond the orbit of Pluto!) rapid convergence was obtained to all three solutions.

It is evident that the new solution procedure for the angles-only problem works extraordinarily well with Herrick's example.

4.2 Examples constructed from Herrick's

Herrick's well-behaved example has been a good basis for the construction of more taxing problems, of which five will be described here. The general basis for construction has been that it is easy to alter the data, and the affine placement of the second observation, in such a way that the first solution of the original problem is still (at least approximately) a solution for the new problem.

In the first constructed example, O_2 was moved to make the new $\underset{\sim}{\lambda}_2$ linearly dependent on $\underset{\sim}{\lambda}_1$ and $\underset{\sim}{\lambda}_3$, since (as noted in section 2.1) this would be enough to invalidate the classical methods. In practice the new $\underset{\sim}{\lambda}_2$ was set to $\frac{1}{2}(\underset{\sim}{\lambda}_1 + \underset{\sim}{\lambda}_3)$, which is not a unit vector but that did not matter. Then an arbitrary position was selected, for O_2, on the affine placement of $\underset{\sim}{\lambda}_2$ required. (The position is 'arbitrary' by the principle of observer independence, the chosen option being to set O_2 in the plane defined by C, O_1 and O_3.) The results obtained are listed in Table 3, without N or ε. Only two solutions could be found, of which the first is the 'correct' (datum) one: it is only slightly less accurate* than the corresponding solution in Table 2, a degradation that is reflected in the condition of the derivative matrix. The second (hyperbolic) solution, though formally 'invalid', is as mathematically acceptable (and accurate) as the first, the obvious question then being "What has happened to the third solution?". The answer is that it has "gone to

* The new solution is only in seven-figure accord with the solution of Table 2, but this is merely due to seven-figure truncation in the coordinates of the new O_2.

infinity", the existence of such a solution being a consequence of the linear dependence*. This was confirmed when the run was repeated with the (exact) eight-decimal components of $\underset{\sim}{\lambda}_2$, which had been constructed to give exact linear dependence, truncated to only seven decimal places: then a third solution emerged with all three ρ_j's of the order 30000 (and only about eight-figure accuracy, associated with severe loss of condition for the derivative matrix) and $c \approx 6 \times 10^{11}$. It was of interest that this solution could be obtained from quite small starting values (<50), without an excessive number of iterations ($N < 20$); in fact, a well-defined value, ρ_{wd}, was found such that, with starting values $\rho_1 = \rho_3 = \rho_0$, for selected values of ρ_0, the first solution would be this special one when $\rho_0 > \rho_{wd}$ but the 'correct' one when $\rho_0 < \rho_{wd}$, without N rising beyond 20 for $\rho_0 \approx \rho_{wd}$.

It being established, from the first constructed example, that the procedure does not suffer when the rank of the $\underset{\sim}{\lambda}$-matrix is only 2, it was natural to construct the second example so as to reduce the rank to 1. This was done by use of parallel lines of sight, with every $\underset{\sim}{\lambda} = (0,0,1)$, and observation sites at $(x_j,y_j,0)$, where (x_j,y_j,z_j) is the location of P_j in the correct solution of the original problem. Four solutions were found without difficulty and are listed, in Table 4, as two pairs: solutions 2 and 4 are mirror images of solutions 1 and 3, the 'mirror' being the xy-plane which is perpendicular to the lines of sight. (No further solutions could be found.) It is easy to see why solutions come in pairs: if an orbit with a given set of elements defines one solution, then another must be the orbit with all elements the same except for Ω and ω which change by π (assuming the inclination, i, to be relative to the xy-plane). In regard to the normal occurrence (when $k = 0$) of solution sets of three, a third pair of solutions might have been expected, but (as in the preceding example) the missing solutions are at infinity; in fact, there are infinitely many solutions at infinity, since the three parallel observations correspond to a single limiting point.

The third example was constructed with $\underset{\sim}{\lambda}_2$ altered so as to be collinear with C, the obvious way to achieve this being to set O_2 at C with $\underset{\sim}{\lambda}_2 = \underset{\sim}{r}_2$: this is a severe version of the circumstance where $\underset{\sim}{\lambda}_2$ lies in the orbital plane, which is prohibited in Escobal's method. The new procedure derived the datum solution without difficulty, but no other solution could be found. Thus there is nothing to tabulate.

In the constructed examples described so far, it will be seen that the actual condition of interest (though not its manner of construction) is unrelated to the solutions, the condition being the rank of the $\underset{\sim}{\lambda}$-matrix or the passage of $\underset{\sim}{\lambda}_2$ through C. In the remaining two

---

* Moving the P_j (for one solution) to infinity is effectively equivalent to moving the O_j towards C, and when the $\underset{\sim}{\lambda}_j$ all pass through C we have the inevitable singularity condition referred to at the end of section 2.1: if (under these circumstances) the $\underset{\sim}{\lambda}_j$ are linearly independent, then there can be no solution; otherwise, the problem is indeterminate in the limit, there being infinitely many solutions.

examples, the condition relates much more specifically to the 'correct' solution of Herrick's example. In the first, we take O_2 at P_2 with an arbitrary direction for $\underset{\sim}{\lambda}$. The objective here was a final verification of the observer-independence principle, in the extreme case of $\rho_2 = 0$. No difficulty arose in the derivation or format of the datum solution (though no other emerged), the associated value of ρ_2 being derived in practice as 4.341×10^{-9}.

The final Herrick-based example to be constructed is of a different nature from the others, the object being to compare the behaviour of the two possible iteration processes when two solutions are so close together that we virtually have a 'repeated root'. The construction was as follows: nearby, or nominal, values $(\rho_{1n},\rho_{3n}) = (2.399,2.824)$ were obtained by truncating the values for the datum solution (Table 2); the corresponding position for P_{2n} was found, for the resulting orbit, via subroutine CALCPS; a transversal P_2P_{2n} to two orbits now being available, a suitable position for O_2 was obtained as the point on this transversal such that $O_2P_{2n} = \rho_{2n} = 2.563$ (arbitrary, but in practice by truncation of Table 2 again) with $\underset{\sim}{\lambda}_2$ defined from $\overrightarrow{P_{2n}P_2}$. Table 5 lists the three solutions for this data to the appropriate accuracy (the third solution, being isolated from the others, is accurate to five more significant figures). Solution 2 is exactly as expected, whilst ρ_1 and ρ_3 for solution 1 are in accord with Table 2; the accord does not extend to ρ_2 because of the way in which the new position for O_2 was obtained.

The number of iterations required for convergence was as follows: with the standard (Halley) iteration process (and the regular starting values), the first solution was obtained in 10 iterations and then the one labelled 3 (from revised starting values) in 7 further iterations, but the one labelled 2 was not reached at all; with the modified Newton-Raphson process (see section 3.3), on the other hand, a single run yielded all three solutions, the number of iterations being 7, 12 and 10, the last solution being the one labelled 2 in Table 5. The relative behaviour for the first two solutions is exactly what one would expect: for 'solution 1', the Halley process required three more iterations, the reason being the proximity of 'solution 2'; then for 'solution 3', derived next, the Halley process took five iterations less, as a consequence of its normal superiority. It is not so obvious why the Halley process failed where the modified Newton-Raphson succeeded, in its final discovery of 'solution 2', on the other hand, since 'solution 1' might now be thought of as being 'out of the way'. The superior behaviour of the modified Newton-Raphson process was confirmed in a number of additional trial runs, however, and is apparently related to a singularity contour in the vicinity of the 'double point'. Under Halley, the 'empirical treatment of overshoot' (described in section 3.3) plays an important rôle, the eventual behaviour being an excruciatingly slow linear progress towards solution 2.

To elaborate on the loss of five significant digits referred to (for solutions 1 and 2) in the last paragraph, this can be attributed to the very poor condition of the derivative matrix in the vicinity of these solutions; thus the two terms of D, in equation (8), then almost cancel,

their values being the same to about four significant figures. In the computation of solution 1, the condition for the matrix is a direct consequence of the proximity of solution 2, the effect being that, with the partial derivatives themselves unusually small, the convergence criterion can be satisfied in spite of the large rounding error in the generation of the final corrections to ρ_1 and ρ_3. The rounding error is unavoidable and there would be no improvement in accuracy from a further iteration.

4.3 Three constructed examples of considerable severity

After the successful results for all the Herrick-based examples, it was decided to construct some severe examples from first principles. The performance of the procedure with three of these examples is now described.

For the most severe example that could be contemplated (and which was soon recognized as officially indeterminate), it was decided to combine two features: the sight-lines $\underset{\sim}{\lambda}_1$ and $\underset{\sim}{\lambda}_3$ would coincide and pass through C; and the positions of the O_j would themselves be compatible with an orbit. The data are listed in Table 6, wherein the first feature (with both observers 'looking in') is immediately apparent; the second feature follows from the evident choice of t_{12} and t_{13} to be $\frac{1}{2}\pi$ and π, respectively, the compatible orbit having $e = 0$ and (with $\mu = 1$) $a = 1$. The sight-line at t_2 lies in the plane through C perpendicular to the coincident lines at t_1 and t_3; more specifically, $\underset{\sim}{\lambda}_2$ is parallel to the z-axis, whilst $\underset{\sim}{\lambda}_1$ and $\underset{\sim}{\lambda}_3$ lie in the y-axis.

Now it may be recognized that there exist two infinite families of solutions to this problem, such that the orbit corresponding to each solution lies in a plane through the y-axis and is symmetric about the xz-plane. In one family the orbits are direct ($i < \frac{1}{2}\pi$) and in the other they are retrograde, the same ellipse being described in opposite directions in a pair of mirror-image paths; the direct family includes the trivial orbit (all $\rho_j = 0$) whilst the retrograde family includes the image of this (with $\rho_1 = \rho_3 = 2$ and $\rho_2 = 0$). In each double orbital plane the point P_2 is uniquely defined by the intersection with $\underset{\sim}{\lambda}_2$, and is the apocentre of the double ellipse for which the values of r_1 and r_3 (equal, and the same for both paths) can be found (and r_2 also) by elementary calculation. (The calculation is most easily based on an assumed value for e: for each e, with $0 \le e \le 1$, r_1 and r_2 may be determined from $\nu_1 = t_{12} = \frac{1}{2}\pi$, via E_1, M_1, n and a.) But not every plane through the y-axis is a possible orbital plane for a pair of orbits, since the limiting cases ($e = 1$) are for inclinations given by $\tan i = \pm\rho_2$, where $\rho_2{}^2 = r_2{}^2 - 1$ and $r_2 = \sqrt[3]{2}$; at this point "i" has lost all meaning, however, the ellipses having become rectilinear and identical. If negative values of ρ_2 are allowed, there are effectively twice as many solutions in each family; they mirror the existing solutions (in the xy-plane) with the mirroring rectilinear ellipse in the limit.

But how does the solution procedure cope with this really severe example? The answer is: surprisingly well, so long as the facility to interchange the rôles of the data at times t_2 and t_3 is utilized! (And the regular starting values could not be used, for a reason that will soon be apparent.) The existence of the two families of solutions with $\rho_1 = \rho_3$ (retaining the original suffix designation for convenience) is demonstrated by convergence to a solution (in the family determined by the starting value for ρ_1) for every choice of a starting value for ρ_2 up to the permitted maximum; ρ_2 remains fixed during convergence, the iteration being effectively on ρ_1 alone. Solutions are found quite fast, and are well-determined unless the fixed value selected for ρ_2 is close to its limiting maximum. (The regular starting values would be above this maximum, which is why they are unusable.)

This behaviour is entirely consistent with expectation. What was *not* anticipated, however, was the appearance of further sets of solutions, in particular of four families of *unsymmetrical* solutions. The characteristic feature of an 'unsymmetric' solution is that the observations at t_1 and t_3 are exactly one revolution apart, so that all these solutions have $a = 2^{-2/3}$; the time intervals t_{12} and t_{23} correspond to differences of π in mean anomaly but to $\frac{1}{2}\pi$ and $\frac{3}{2}\pi$ (or vice versa) in true anomaly. The existence of four families, of essentially equivalent orbits, results from the combination of the two possibilities for true-anomaly differences with the possibilities for both direct and retrograde orbits. The orbits in the xy-plane have the value $e = 0.617...$ (as opposed to zero for the symmetric orbits), but at the other extreme (maximum inclination) the limit is provided by the same rectilinear ellipse as constitutes the limit for symmetric solutions. Since this rectilinear orbit is the limit of no less than six families of solutions, two symmetric and four unsymmetric, it is not surprising that the procedure breaks down in its vicinity.

The unsymmetric orbits, like the symmetric ones, are obtained with ρ_2 remaining constant through all the iterations. Since k, the half-revolution count, is based on t_{12} rather than t_{13} (because of the rôle interchange), half the unsymmetric solutions are found with $k = 0$ whilst the other half require $k = 1$. But with $k = 1$, two further families of symmetric solutions appeared; the relation of these to the original two families is that the swept-out angles corresponding to t_{12} and t_{13} are $1\frac{1}{2}\pi$ and 3π, not $\frac{1}{2}\pi$ and π.

The final total for the number of solution families is now seen to be eight. To illustrate, Table 7 lists all the eight solutions in the xy-plane, to the appropriate accuracy in each case, the values of ρ_2 (all zero) being omitted but with k and family-type included.

The second constructed example was a distinctly less severe version of the first, the differences being as follows: the positions of O_1 and O_3 were moved closer to O_2, to remove the collinearity with C, the adopted coordinates being $(0.6,\mp 0.8,0.1)$; the corresponding sight-lines were given non-zero z-components, being (arbitrarily) specified as $(0,\pm 1,0.5)$; and the times were reduced in conformity with retention of the trivial circular

orbit as a solution, the new value of t_{12} being given by arc-tan (4/3). The problem was now fully determinate, solution becoming possible without the interchange of data points. Only two solutions could be found, however, the non-trivial one being $\rho_1 = \rho_3 = 0.43732746883$ and $\rho_2 = 0.32596464957$, with $e \approx 0.71$ and $i \approx 18$ degrees. So there was little to be learned from this example.

It was noted with the example of extreme severity that solutions could not be obtained near the pair of rectilinear orbits that are multiple limits corresponding to extreme values, positive and negative, of ρ_2. This was hardly surprising. But it suggested an arbitrary rectilinear orbit as the basis for a final constructed example. Assuming such an orbit along the positive z-axis, with both a and e set to unity, the values of E at P_1, P_2 and P_3 were taken to be $\frac{1}{3}\pi$, $\frac{2}{3}\pi$ and π respectively, from which the values of $z\ (= r)$ are 0.5, 1.5 and 2.0. The resulting values of t_{12} and t_{13} are $\frac{1}{3}\pi$ and $\frac{2}{3}\pi + \sqrt{3}/2$, as included in Table 8 (but hand-calculated to 10 digits only); the other tabulated entries follow from the setting of O_1 on the x-axis, O_2 on the y-axis and O_3 to coincide with C.

Run (with default starting values) in the usual way, the correct solution was obtained in four iterations, and is the first of those given in Table 9. The values obtained for ρ_1, ρ_2 and ρ_3 are in proper accord (to the accuracy of the data) with the correct values of $\sqrt{1.25}$, $\sqrt{3.25}$ and 2. The correct values of a (1.0), e (1.0) and i(90°) were obtained, with conventional values of Ω and ω since an orbital plane is not defined for a rectilinear orbit. Solution 2 is a second rectilinear one that emerged, together with the other six solutions (non-rectilinear ellipses), when runs in the interchange mode were undertaken. The value of ρ_1 is the same in all eight solutions because P_1 (as well as P_3) always lies on the z-axis. The non-rectilinear orbits involve a true-anomaly arc coverage (from P_1 to P_3) of 2π (solutions 3-6), π (solution 7) or 3π (solution 8); these are basic-singularity conditions, which is why solutions 3-8 can only be found in interchange mode. The overall conclusion, for the rectilinear-orbit example, is that the behaviour of the procedure with all its subordinate subroutines and functions[2,3,14], is remarkably good!

4.4 Two examples of Escobal

In the course of an extended correspondence between Mark Lane of the MIT Lincoln Laboratory and the present author, Lane drew attention to five problems that he had taken from the examples and exercises in Chapter 7 of Escobal's textbook[10]. Three of these could be solved without difficulty by Lane's own method[13], but something was amiss with the other two and the author was invited to apply the new method. As a result, it was possible to confirm Lane's conclusion that a major error (such as the date or times) must be present in the data for the first problem, but that if allowance is made for the effects of perturbations etc in the second example it might be possible to reconcile the textbook solution with the (only) one found by Lane and the author. Unlike Herrick, Escobal supplies only the final

answers, as sets of orbital elements, so no inferences are available from intermediate stages of the computation.

The first problem is Exercise 1 of Escobal (*loc cit*),which relates to the Earth satellite 1959 Alpha 2. If we express distances in kilometres and times in seconds, with the value $\mu = 398601.3$, then the data for the three observations (ignoring light-time corrections and all other complications) are as in Table 10. The five solutions obtained (in the usual way, except that the fifth solution only emerged during a later run) are as listed in Table 11, and it is seen that no solution can be accepted -- in particular because there is at least one negative range in each case. To avoid the possibility of error in the data preparation, the program was also run with the data that had been prepared independently by Lane. The units for Lane's data were Earth radii for distance and days for time (with $\mu = 11467.89807$) but essentially the same results were obtained.

In view of his failure to obtain a valid solution at all, Lane computed the sight-lines that would be needed for consistency with Escobal's nominal solution, assuming no other change in the data. A complete set of 'revised data' is presented in Table 12, therefore, exactly as supplied by Lane in his own preferred units and obviously entirely different from the data in Table 10. The solutions are given in Table 13, where it is Solution 3 that agrees with the nominal result. The notable feature of these results is the existence of a fourth solution when $k = 0$. This was entirely unexpected and it was gratifying that the four solutions were found in the usual way (with revision of starting values) without difficulty.

The second problem was Exercise 6 of Escobal (*loc cit*), which relates to the Earth satellite 1959 Eta. Table 14 gives the data (units kilometres and seconds) and it will be seen that this is a very-short-arc problem with all the observations from the same station. The only solution that could be found ($k = 0$ of course) is given in Table 15. This is formally accurate to the number of figures quoted and the value of the eccentricity (0.197) agrees with that obtained by Lane, though Escobal's official solution involves $e = 0.232$. The result is obviously very sensitive to any allowance for noise in the data, however, and to the inclusion (or otherwise, as here) of the effects of orbital perturbations, light-time corrections in the data, etc. So it may be concluded that there is not necessarily any real discrepancy present for this example.

4.5 A general survey of single-revolution examples

In addition to the real-life and artificially-constructed examples already described, it was natural to wish to make a grid-type survey of all possible nominal orbits (hyperbolic as well as elliptic), with a similar survey of all possible sets of three possible observations at all possible times. But far too many parameters are involved for all grid points to be covered without the performance of billions of test runs, so for the exercise to be reduced to manageable proportions the coverage had to be severely restricted. In practice the following

restrictions were introduced (with $\mu = 1$ in all cases): fixed positions of the observing sites were selected; times would always be at uniform intervals (so that $t_{12} = \frac{1}{2}t_{13}$); a particular orbit would serve as the 'basic nominal'; and a variety of test cases would be generated via families in each of which a single orbital element, or else the value of t_{13}, would be varied. Each test run consisted of the generation of the observations corresponding to the nominal orbit under current consideration, followed by the derivation of all the solutions that the new procedure (run without prejudice of course) could locate.

On an arbitrary basis the observing sites were placed, once and for all, at (0.1,0,0), (0,0.1,0) and (0,0,0.1), and the basic nominal orbit was defined by $(a,e,i,\Omega,\omega,\tau) = (1,0,0,0,0,0)$, implying an orbital period given by $T = 2\pi$; in the family of e-varying runs, a was changed to -1 for the (hyperbolic) cases with $e > 1$. The basic value of t_{13} was taken as 4, whence $t_{12} = 2$ and (since $t_{12} < \frac{1}{2}T < t_{13}$ for elliptic orbits) all examples are indeed single-revolution but with $k = 1$ not 0 (for ellipses). Again, since CALCPS assumes the time of the first observation to be the time origin (that is, $t_1 = 0$), the implication of $\tau = 0$ is that the first observation is made at pericentre, itself a severe test of the procedure for highly eccentric orbits; a further implication is that the second observation occurs (for elliptic orbits) before the next apocentre, so that if the facility to interchange the second and third observations was invoked, as it always was if the 'true' orbit could not be recovered under the normal mode of operation, then it would be necessary to switch to $k = 0$.

Results from the successive families of test runs will now be given.

(a) Examples for a range of inclinations

In the first family of test runs, seven values of i were used, viz 0°, 30°, 60°, 90°, 120°, 150° and 180°, the first of these defining the basic orbit itself. No difficulty was encountered in the recovery of the nominal solution in each case, in at most five iterations, with at least one additional solution always found. For $i = 0°$, 60° and 180°, a third solution was found, whilst for $i = 30°$ a total of four emerged.

With $\Omega = 0°$, it is clear that both O_1 and O_3 lie in the true orbital plane when $i = 90°$, whilst O_1 and O_2 lie in the orbital plane when $i = 0°$ and 180°. Thus these may all be regarded as quite severe tests. It will be seen, in the next section, that a non-circular orbit can be a source of particular difficulty when $i = 90°$.

(b) Examples for a range of eccentricity

It was decided to run with the following values of eccentricity (input, for convenience, rather than the 'universal' element, q): 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.999, 1.001, 1.1, 2, 4, 10 and 100, the elliptic orbits being for $a = 1$ and the hyperbolic orbits for $a = -1$ as already noted. It was assumed (and confirmed) that a run

with $e = 1$ (rectilinear ellipse or hyperbola, according to the sign of a) would fail, in particular because the first observation would then be of the force-centre itself! By the same token, the orbits with $e = 0.999$ and 1.001 would then furnish tests of considerable severity.

The initial intention to operate with a single family, using the basic value of $i = 0°$ for all runs, was revised in favour of three families, using the additional values $i = 45°$ and 90°. For $i = 0°$, the correct solution was recovered for all orbits other than with $e = 0.999$, at most one other solution being obtained. At most nine iterations were needed for the elliptic orbits, not counting the complete failure ($e = 0.999$), and five sufficed for each of $e = 2$, 4 and 10, but the other hyperbolic orbits required more iterations, the number being 20 for $e = 100$. Under normal operation, no solution was found at all for $e = 0.999$, but two were found in the interchange mode; unfortunately neither corresponded to the nominal orbit.

The results for $i = 45°$ were in some ways (but not all) better, as expected. Far more solutions were found for the elliptic orbits, including the correct one in all cases but (again) $e = 0.999$; in fact three solutions were found in all cases (up to $e = 0.9$) except for $e = 0.6$ (two solutions) and $e = 0.7$ (four!). More iterations were required, however, rising to 15 for $e = 0.8$. As with $i = 0°$, the correct solution was found for all the hyperbolic orbits, but the interchange mode was needed for $e = 4$ and 10. The recalcitrant case ($e = 0.999$) was again not helped by interchange.

The results for $i = 90°$ were (as expected) the worst of the three sets, though for hyperbolic orbits there was little difference from the other two sets. For elliptic orbits the correct solution was only found with $e = 0.1$, 0.3, 0.4, 0.5 and 0.6, the full number of iterations (25) being required for $e = 0.5$, and the interchange mode did not help any of the other cases. The failure for $e = 0.2$ was entirely unexpected and, upon investigation, revealed the behaviour that was remarked upon in section 3.3, such that with the regular starting values, $(\rho_1,\rho_3) = (0.6,0.6)$, the unmodified Halley process 'converged' with ever-increasing slowness to the point (-0.2220...,0.1604...), which is not a solution of the problem but at which the derivative matrix is totally singular. Apart from the true solution, $(\rho_1,\rho_3) = (0.7,1.2089...)$, two other solutions were found by experimentation, (0.4046...,-0.6457...) and (-1.4348...,0.9591), and the false-convergence phenomenon should be explicable in terms of the configuration of the three solutions; it is naturally much more complicated than the simple one-variable quadratic equation considered in section 3.3, however. The effect of the modification (involving geometric means and described in section 3.3) was somewhat paradoxical. On the one hand, there was still no convergence (within 25 iterations) to any solution when the regular starting values were used. The problem of 'decelerating convergence to a singular point' disappeared, on the other hand;

moreover, when starting values close (by four-figure truncation) to the singular point were used, a circuitous convergence to the right solution was achieved after just 25 iterations.

Particular attention was paid to the case just described because difficulty had not been expected at such a low eccentricity and because the investigation was revealing. The modification to the iteration process was certainly worth making, but the difficulty (in routine operation) for this case was not eliminated; the reason for this requires further investigation.

(c) Examples with alternative values of Ω, ω and τ

To avoid proliferation, it seemed sensible to use only one alternative value for each of the elements Ω, ω and τ, the full range of values of e being covered to provide a family of runs in each case. It was decided to use the value $i = 45°$, rather than $i = 0°$, with each family.

With the alternative value $\Omega = 45°$, it was no longer the case that the first observing site (O_1) was automatically within the true orbital plane. The results were very little different from the family with $\Omega = 0°$, however. Two cases failed: for $e = 0.999$, as before; and for $e = 1.001$. But the latter was not significant, as there had only just been convergence (24 iterations) for $\Omega = 0°$.

For the next family, the alternative value $\omega = 45°$ was introduced, Ω being retained also at 45°. The results were, unfortunately, marginally worse, since the correct solution could now not be obtained for $e = 0.6$ and 0.8 as well as 0.999. A remarkable feature of the run with $e = 0.9$ was the derivation of a record number of solutions, six, when operating in interchange mode with $k = 0$. For the hyperbolic orbits, however, a single solution, the correct one, was obtained for every value of e.

For the final family, τ was set to the value corresponding to a mean anomaly (elliptic or hyperbolic) of 45°, ω being restored to zero and Ω retained at 45°. This meant that the use of pericentre as the first observed point was at last being abandoned, the gratifying effect being the virtual disappearance of all convergence problems, the only failure being for the largest value (100) used for e. This shift of the first observation away from pericentre meant, however, that it was no longer appropriate to set $k = 1$ for all the elliptic runs and $k = 0$ for all the hyperbolic runs; the switch from $k = 1$ was now required between $e = 0.2$ and $e = 0.3$. At most five iterations were needed to obtain the correct solution for any elliptic orbit, including (for the first time) the rectilinear ellipse. For the hyperbolic orbits, the maximum number of iterations needed was ten until $e = 10$, at which point 19 were necessary. It has already been noted that convergence failed for $e = 100$, no solution at all being found with the usual starting values, (0.6,0.6), or even with (0.69,0.69) come to that; with starting values (0.7,0.7), however, the correct solution was found in twelve iterations. So even the failure for $e = 100$ was marginal.

(d) Examples with a range of values for a and t_{13}

After the runs already described, it remained to consider the scale effects, in distance and time, corresponding to a range of values of a and t_{13}, with the value $t_{12} = \frac{1}{2}t_{13}$ maintained. The values of the other orbital elements were fixed as for the last family of runs described (alternative value for τ), just a limited range of eccentricity being covered.

The range of semi-major axis covered was, *ipso facto*, different as between elliptic and hyperbolic orbits. Thus, for two 'spot' values of $e < 1$, viz $e = 0.2$ and $e = 0.9$, the families of runs were for a = 1, 1.2, 1.5, 2, 3, 5, 10, 20 50, 100, 200, 500 and 1000 (values of $a < (2/\pi)^{2/3}$ would be inappropriate for nominally single-revolution orbits). The correct orbit (with no alternatives) was derived, usually in six iterations or less, for all combinations of (a,e) with the single exception of (1000,0.2). For hyperbolic orbits, a single value of e was adopted, viz $e = 10$, the coverage of semi-major axis being the negatives of the values for elliptic orbits, supplemented by the values a = -0.1, -0.5, -0.7, -0.9. The results now were rather poor, since the correct orbit was only derived for the values a = -0.7, -1, -2, -10 and -100, with no solution at all for the other cases.

For the family of runs with t_{13} varying, a single orbit was used, with $a = 1$ and $e = 0.1$. It will be recalled that the basic value was $t_{13} = 4$, this being the mean-anomaly difference in radians. For increasing values of t_{13} up to the limit of 2π, corresponding to multirevolution coverage, it was expected that difficulties would be encountered as the limit was approached, in particular because of the Lambert indeterminacy that would arise with an angular separation approaching 2π (or π when running with interchanged observations) as well as one of π. In practice all was well up to $t_{13} = 6.1$, but at $t_{13} = 6.2$ the problem was met, aggravated (it would appear) by the fact that no less than six solutions were eventually found after experimentation with this particular case. With *decreasing* values of t_{13}, on the other hand, entirely predictable results were obtained. They may be summarized by the remark that the use of τ as one of the author's 'universal' elements[14] involves the addition of τ to t_{12} in CALCPS, so that the accuracy of the overall solution process inevitably degrades without limit as t_{12} and t_{13} decrease in magnitude relative to τ.

4.6 A range of multirevolution examples

It was decided not to undertake a full (general) survey of multirevolution problems, but to restrict consideration to a single family of runs based on the particular orbit specified by (a,e,i,Ω,ω,M) = (1,0.1,45°,45°,0°,45°). The family of runs is an extension of the last one used in section 4.5, with t_{13} changing from run to run (and $t_{12} = \frac{1}{2}t_{13}$), the values of t_{13} starting where they previously left off. The single-revolution results indicate that it is only in regard to e that runs for other values of the orbital parameters might have been significant, but for multirevolution orbits (with $e < 1$, necessarily) it seemed unlikely that runs with $e > 0.1$ would be of great practical interest.

In regard to the values of t_{13} selected, it is remarked that the value $t_{13} = 2j\pi$ (for any integer j) corresponds to the spanning of exactly j revolutions between the first and last observations. It was convenient, therefore, to use integral values of t_{13}, starting with $t_{13} = 7$, to introduce a quasi-random variation in θ, the reduced angle P_1CP_3, as t_{13} was increased. To cover nearly eight revolutions, as an arbitrary limit, all (integral) values of t_{13} up to 24 were selected and then alternate values (even integers) up to $t_{13} = 48$. Thus the total number of runs was 30.

The regular procedure, with standard starting values, recovered the correct solution without difficulty (without the restart with common-perpendicular starting values coming into play, in particular) in 24 of the 30 runs, the maximum number of iterations required being six. The correct value of k had to be supplied for each example, of course (several values might have to be guessed in turn, in real-life runs), and the correct solution could correspond to either the higher-energy Lambert assumption or the lower-energy assumption. Additional solutions were found for most of the examples, and these did sometimes involve the restart procedure (which comes into play automatically). The failures for the remaining six examples could all be attributed to, or at least associated with, insufficiently close starting values, since it was checked in each case that convergence followed rapidly from starting values close to the known solution; this would be of no comfort in real-life problems, of course, unless good estimates happened to be available. For three of the six failing examples, the solution could be found by use of the facility for interchange of the second and third data sets, but interchange did not help the other three.

To be specific, the failures were for $t_{13} = 9$, 15, 16, 22, 28 and 44. An immediate conclusion is that there is no marked tendency for a failure to become more likely as the number of included revolutions increases; this is one reason why the number of included revolutions was not taken beyond eight. The simplest failure to explain is the one for $t_{13} = 44$, since this example involved an angle of only 1.15° for θ. The unreduced angle (for the correct orbit) includes seven revolutions, its value being 2521.15°, with $k = 14$; the root of the problem for this value of t_{13} was the proximity of π to 22/7. Somewhat surprisingly, this was one of the three examples for which the solution could be found by the interchange facility, the other two being for $t_{13} = 15$ and $t_{13} = 28$.

4.7 The Molniya example of Lane

The final example to be presented is the second of the three presented in the recent paper by Lane[13] that has already been referred to. The example was based on a Molniya-type orbit for which the nominal elements at epoch are given by (a,e,i,Ω,ω,M) = (4.16347314,0.74,63°,200°,280°,300.541°), where the semi-major axis is in Earth radii, the value of μ being taken as 11468 in the appropriate units. The data (taken directly from Ref 13) are listed in Table 16.

An attempt was made to find all possible solutions for this problem, the results being presented in Table 17. In running the new method in the standard way described in section 3, with revised starting values (as at the end of section 3.5) for multiple solutions, 193 solutions were found, as indicated by the index numbers (when present) in column 1 of the table. Of these, 137 can be dismissed as impossible, either because at least one of the three ranges (in columns 4-6) are negative or because the perigee, at geocentric distance $a(1 - e)$, with approximate values of a and e given (in columns 7 and 8) in the table, is subterranean. The impossible solutions are indicated by the asterisk in (column 2 of) the table, the 'possible' solutions that were found automatically (and rapidly) being the rest of the numbered ones. The solutions were obtained by running with successive values assumed for k (the number of half-revolutions spanned), the value of which is included (column 3) in the table. As no solution could be found with $k > 159$, with $a \approx 0.961$ (about 6129 km, corresponding to $T = 1.327$ h), it may be assumed that this is the energy-limiting value of k referred to at the beginning of section 3.6.

Lane himself tabulated only eight possible solutions to his Molniya problem numbered 2, 5, 6, 8, 9, 10,11 and 12 in Ref 13; his procedure suppressed solutions with any range negative, so the gaps in his numbering sequence were entirely due to subterranean perigee. It had been assumed that most of Lane's solutions would be included in the 56 'possible' ones derived by the new method in standard mode, but in fact only his Solution 12 emerged – as the solution numbered 13 in Table 17. The other seven were obtained in the course of non-routine operation of the method, however, and are included, with others derived in the same way (mostly by the use of unchanged starting values in deriving multiple solutions), in the table; such solutions, unnumbered in column 1, appear after the numbered ones, except that the unnumbered solutions of high energy (for a given value of k) are placed before the numbered solutions of low energy. Lane's solutions may be identified by the presence of his solution number in column 2 here, since an asterisk cannot be needed for these solutions. The additional solutions derived in this way were 62 in total, of which only eight (the seven remaining Lane solutions and a solution with $k = 19$) are possible. Thus a grand total of 255 solutions were found, with 64 deemed possible. The correct solution is Solution 5 of Lane, between the regular solutions numbered 45 and 46 here (the value of a in this solution corresponds to $T = 11.96$ h).

It may be thought that the new method's performance is curiously unsatisfactory for Lane's example, since (in routine operation) it located only one of the eight Molniya-type orbits that Lane found as solutions, and this one not the 'correct' one. (It is worth remarking that four of Lane's solutions are 'direct' orbits with $i \approx 63°$ and four are 'retrograde' with $i \approx 117°$; for the full set of 255 solutions it is similarly true that about half the orbits are direct and half retrograde, but the ranges of inclination covered are not so closely centred on 63° and 117°.) These failures are highly correlated, however, since in

each case we see from the table that (for a given k) Lane's solution is preceded by a solution of lower eccentricity but almost identical semi-major axis (and inclination likewise, it happens) with both ρ_1 and ρ_3 close to 6. Now the starting values for the first solution are (in all cases) given by $\rho_1 = \rho_3 = 6$ (and would be exactly 6 if the observing sites were all at exactly one 'Earth radius' from the centre of the Earth), so the solution procedure starts (for each value of k for which there is a 'Lane solution') by finding this lower-e solution quite rapidly. After this it appears that (with one exception) the multiple-solution algorithm is insufficiently robust to locate Lane's solution, the precise reason for this being unknown.

It will be noted that for all the lower-e solutions referred to in the last paragraph (with ρ_1 and ρ_3 both near 6) the value of ρ_2 is negative; in fact, the difference between the values of ρ_2 for the possible and impossible solutions is always much greater than the corresponding differences for ρ_1 and ρ_3. Two points may be inferred from this. First, a certain amount of bad luck was associated with the failure to locate the possible solutions, in particular because of the solution procedure's use of ρ_1 and ρ_3, but not the more sensitive ρ_2, as basic unknown variables. Second, there might appear to be some advantage in a preliminary 'starter modification process', prior to the main iteration process, along the following lines: from the procedure's standard value of ρ, defined as the common starting value of ρ_1 and ρ_3, we obtain the corresponding value of ρ_2 in the normal way; unless this is negative, we start the main process as usual; if $\rho_2 < 0$, on the other hand, we compute a numerical value for $\partial\rho_2/\partial\rho$ (just as for the derivatives f_x etc in section 3.3, our concern now being with the unknown range along the observed line of sight, rather than the known direction); from ρ_2 (< 0) and $\partial\rho_2/\partial\rho$ we obtain a revised value of ρ, corresponding (on a linearity assumption) to $\rho_2 = 0$. This 'starter modification' could be enhanced, if desired, by iteration or by use of the second derivative $\partial^2\rho_2/\partial\rho^2$.

The objective of starter modification is to encourage the solution procedure to go directly to a valid solution, rather than via an invalid one, in particular for problems such as Lane's. In the basic form outlined, the modification was so simple to implement that it was applied to Lane's example after the first draft of this Report had been completed. The results were disappointing, however, and it would appear that a more subtle version of the modification is required. It must be admitted, in the meantime, that the current procedure does not perform as well, for Lane's example, as had been expected from its excellent performance for the previous examples.

5 DISCUSSION

Previous researchers[10,12,13] have assumed that, when ρ_1 and ρ_3 are the iteration variables, it is logical to take t_{12} and t_{13} (or t_{23}) as the target functions. This certainly leads to rapid computation, since there is no need to solve Kepler's equation, let alone the

Lambert problem, but the speed is achieved at the expense of a severe loss of general applicability (robustness), as suggested in section 3. To confirm this suggestion, it was decided to program the usual, or 'standard', procedure, to compare its performance with that of the new procedure. The comparison was confined to three 'real world' examples, as provided by Herrick, Escobal and Lane, since it was already known that the standard procedure would fail with severely taxing artificial examples.

Before the results of the comparisons are described, it is worth indicating why the standard approach is prone to such poor performance unless the starting estimates are good – and sometimes even then. As can be seen from Appendix A, the basic distinction between the two approaches is as follows: after the current estimates of the ranges ρ_1 and ρ_3 have been used (in both approaches) to define the (hypothetical) orbital plane and two positions within it, the new procedure introduces the times, and hence the dynamics, at once; the standard procedure, on the other hand, continues with geometry alone, obtaining the implied additional position within the plane; then the parameters of the orbital path are determined, and only after this are the dynamics brought in, computed time intervals being derived for comparison with the observed time intervals (as opposed to a computed line of sight at time t_2, for comparison with the observed one). The critical parameter is p, the semi-latus rectum, the (purely geometric) formula for which is immediate from the three linear equations (A-1) in Appendix A. To be dynamically acceptable, the value of p must be positive (for parabolas and hyperbolas as well as ellipses), so if $p \leq 0$ the standard approach must fail though the new one does not. (Negative p is *geometrically* acceptable but corresponds to the 'dynamically wrong' branch of a hyperbola that may be interpreted with $a > 0$ and $e < -1$; this branch is unacceptable because it implies that the inverse-square-law force is a repulsion!)

If we consider a continuous evolution of estimates for ρ_1 and ρ_3, in a given angles-only problem, there are three ways in which p can turn negative: first, and most obviously, the denominator (in the formula for p) can pass through zero, at which event the vectors $\underset{\sim}{r}_j$ represent three collinear points, the implied transition from one branch of a hyperbola to the other being evident; second, the numerator can pass through zero, at which event two of the three $\underset{\sim}{r}_j$ point in the same direction; third, the (transitional) orbital plane can be parallel to the line of sight corresponding to $\underset{\sim}{r}_2$, so that this vector is of infinite magnitude and instantaneously reverses its direction. In this last circumstance, neither the numerator nor the denominator of p actually passes through zero, but the latter reverses sign while the former changes its value discontinuously without a change of sign (if the numerator changes its sign as well as the denominator, then p obviously retains its sign); thus the limiting legitimate orbit in this case (unlike the other two) is not a degenerate one! (To appreciate the significance of the right choice of target functions in iteration procedures generally, it is instructive to consider the relatively elementary problem of solving Kepler's

equation, $E - e \sin E = M$, with M in the range $[0,\pi]$. Then M is of course the 'right' target function, with e treated as a pure constant, and[1] a possible universal starting estimate for E is given by $E_0 = \pi$; if the rôles of M and e are reversed, on the other hand, then to set $E_0 = \pi$ would be fatal, since it leads to infinity as the computed value of e.)

To make comparisons for the three examples referred to, we look at what happens with various starting values, restricting ourselves to choices with $\rho_1 = \rho_3$, both equal to ρ, say. For Herrick's example, we find that most positive values of ρ lead to negative p; ρ has to be less than about 2.96, or else lie within the range [3.16,3.49], for p to be positive. But the wanted solution is given by $(\rho_1,\rho_3) = (2.39...,2.82...)$, so we can see how small a margin there is between this solution and starting estimates that must fail at once – contrast this with the performance of the new procedure, which (as noted in section 4.1) converges rapidly to all three solutions, even from ludicrously high values of ρ. It was found, in fact, that from no value of ρ at all could the wanted solution be obtained, though the other two solutions could both be obtained (by the multiple-solution facility, carried over from the new procedure to the standard one) from $\rho = 2.0$, in particular.

The example of Escobal used for comparison purposes was the second one from section 4.4, in view of the unresolved problems with the data for the first example. In this case there is no negative-p region near the solution and the standard method succeeds with quite large values for ρ (as starting estimates for ρ_1 and ρ_3), though not for the virtually unlimited value that is permissible with the new method. Over the complete range $[-\infty,\infty]$ for ρ, p changes sign six times, including all the three ways that have been described.

With Lane's example, an important aspect of the comparison arises from the fact that multiple revolutions are involved. For single-revolution problems, we have seen how much more robust the new procedure is because the subordinate Lambert problem always has a solution, whereas the subordinate Gibbs problem, in the standard procedure, does not. For multiple-revolution problems, on the other hand, the new method is adversely affected by the fact that the Lambert problem now does not always have a solution; but the Gibbs problem is no less tractable than in the single-revolution case, since the derivation of p is independent of any assumption involving the number of included revolutions. Hence much of the advantage of the new method might be expected to be lost when the true solution of Lane's problem (the one labelled 5 in column 2 of Table 17) is sought by both procedures. Now there are actually five solutions with $k = 17$ in Table 17, of which four involve positive ρ_1 and ρ_3 that do not differ very much from solution to solution, so we simplify the comparison by looking for any one solution (with $k = 17$) by the two methods, with starting estimates given by $\rho_1 = \rho_3 = \rho$ for various ρ. The standard method, which gives positive p for starting estimates with ρ in the range [4.5,15.1], only yields a solution if ρ is in the range [5.0,7.3]. The new method, on the other hand, delivers a solution for all

positive ρ up to about 6.7, without advantage being taken of the restarting techniques described in section 3.6, and for unlimited ρ if advantage is taken. So the new method is likely to be more robust even for multirevolution examples.

To conclude this discussion, a curious finding with Lane's example is worth recording. Whilst looking at regions of negative p in the vicinity of actual solutions for (ρ_1,ρ_3), it was found that the first solution with $k = 1$ (the one labelled '2' in column 1 of Table 17) was extraordinarily close to such a region. Thus, the solution is [6.226..., 5.641...], yet the rounding of this to two significant figures produces starting estimates that must fail, since $(\rho_1,\rho_3) = (6.2,5.6)$ lies in a region of negative p, the boundary of which is marked by transitions of the third type above.

6 CONCLUSIONS

The ultimate goal in the development of algorithms for solution of the minimal angles-only orbit determination problem may be summarized as follows: an algorithm is sought that is capable of locating all solutions of an arbitrary 'general problem'; solutions should be determined in the shortest possible time on the computing device that is used, and to the maximum possible accuracy inherent in the data; intrinsic limitations, associated with some form of indeterminacy, should be recognizable; and there should be options to increase efficiency by the avoidance of solutions that are incompatible with any legitimate assumptions about the orbit. The traditional approaches, developed for distant heavenly bodies and slow hand-computing methods, are not usually appropriate for artificial satellites and modern computers, so recent years have seen the development of new approaches. The approach described in the present Report is based on the author's algorithm for Lambert's problem and has led to a solution procedure with several novel features.

The procedure is based on the iteration of values for two of the three unknown ranges, with novelty from the placing of a Lambert solution algorithm at the heart of the procedure. The corollary of this approach is that solutions can, in principle, always be obtained to the accuracy inherent in the data for the problem, with no difficulties associated with the artificial singularities that have been responsible for failure conditions with other methods. The reason for the qualification 'in principle' is not that solution sometimes fails when iteration is based on the first and last observation, since it is a trivial matter to switch to a different pair; the qualification is necessary only because the question of suitable starting values has not been fully resolved. With many problems, all solutions can be obtained from almost any starters; other problems are much less tolerant, however.

The approach to multiple solutions is believed to be completely new, its essence being the modification of the iteration process in such a way that solutions already treated are not found again. There is unfortunately no guarantee, however, that every solution to a

given problem will indeed be found. The difficulty is again associated with the starting values, and two different formulae for assigning these have been tried; the first is just to use the same values for all solutions, whilst the second utilizes the values to the solutions previously obtained.

It follows, from the critical use made of the author's Lambert algorithm in the new angles-only procedure, that the procedure must be told how many half-revolutions to include between the observations whose ranges are iterated. It will often be necessary, in locating the right solution, to try more than one value of this integer, k, the solutions being single-revolution if $k = 0$ or 1, but multirevolution if $k \geq 2$. Thus the coverage of multi-revolution ellipses is inherent in the approach, and this also is believed to be novel. New difficulties arise with multirevolution problems, however, in particular because the Lambert problem does not always have a solution when coverage exceeding one revolution is assumed.

The procedure has performed extremely well with a number of test problems, notably with some that were expressly set up to test to the limit. With other problems it was less successful, usually those involving multiple iterations and in particular the example taken from Lane's study. The good news with Lane's example was that the procedure, running in an automatic mode (with increasing values of k), was able to find 193 solutions to the problem; the bad news was that it missed at least another 62, including the only solution that was really worth having – being the 'correct' one.

Several ways forward are possible with the present form of the new angles-only procedure (as listed in an Appendix) as starting point. It could be made more efficient for operational use; in particular, the partial derivatives could be derived analytically, rather than numerically. The choice of starting values could be changed to take more account of the knowledge of the 'real world' for a particular application, perhaps via the input of estimated values for one or more of the orbital elements. And the use of additional observations could be permitted. With this last enhancement, however, the problem would no longer have the minimal nature that was postulated *ab initio*.

Acknowledgment

The author was greatly stimulated by the paper[13] by Dr Mark Lane (of the MIT Lincoln Laboratory), originally presented at the 1991 Astrodynamics Conference, held in Durango, Colorado. The subsequent correspondence with Dr Lane was very helpful.

Appendix A

THE THREE-VECTOR PROBLEM OF GIBBS

We may describe Lambert's classical problem as a two-vector one, since it consists in the determination of an orbit from vector positions at two given times. This makes it a minimal problem, since the six vector components are just what are necessary for the derivation of six orbital elements. Thus the problem has always at least two solutions[3], the total number being dependent on the interval between the two times. It is far from a trivial problem, however, and some of the difficulty is a direct consequence of its minimal character.

It was recognized by Gibbs that things are much simpler when a third vector is available, since in principle we now have enough data to ignore the times, just as in the angles-only problem we do not need observations of range, which are assumed not to be available. It is not quite as simple as that, however, a self-consistent problem being over-determined even without the times, since three arbitrary vectors would not be compatible with the Keplerian assumption of coplanarity. Thus only five elements can be determined when the times are ignored; but this is as it should be, of course, since a time item is needed to distinguish between the 'orbits' that share a common 'path' (via the mean anomaly, or whatever sixth element is in use) – the other two times, and likewise one (or more) of the ranges, can be used to check the validity of a computed solution.

In regard to the number of solutions to the problem with five items of data, the general situation is that there are exactly two, involving the same orbital path but described in opposite directions; only one of these is compatible with the times. The multirevolution problem, which so bedevils the Lambert and angles-only problems, does not arise, since the times (if compatible with one of the only two possible solutions) will automatically reflect multirevolution arcs.

Our data, then, comprise the three vectors, $\underset{\sim}{r}_j$, and the first step is to specify the orbital plane. This may be done, in general, via $\underset{\sim}{r}_j \wedge \underset{\sim}{r}_k$ for any two of the three vectors, since each product gives a vector normal to the orbital plane. This assumes consistency, with a consistency check implicit, and for maximum accuracy it is natural to select the vector product of greatest magnitude. The fact that interchange of $\underset{\sim}{r}_j$ and $\underset{\sim}{r}_k$ reverses the sign of the original vector product reflects the fact that there are two 'orbital' planes, representing motion in opposite directions in the same 'geometric' plane. If one of these is described (relative to the axis system in use) by the elements i and Ω, then the other is described by $\pi - i$ and $\Omega + \pi$. If $u\ (= \nu + \omega)$ is the angle such that motion in the orbital plane is described by the polar coordinates r and u, then u for the one plane is associated with $\pi - u$ for the other. (For a discussion of the transformations to and from the coordinate

system based on the orbital plane, including the avoidance of all difficulty associated with indeterminacy or singularity, see Ref 14.)

It remains to determine three orbital elements (say, a, e and ω, but again see Ref 14) from the three in-plane positions (r_j,u_j), but this reduces to the solution of simultaneous equations in three unknowns. Thus the equation

$$\frac{p}{r} = 1 + e\cos\nu \qquad \text{(A-1)}$$

holds for any non-rectilinear orbit, and is linear in the unknowns p, $e\cos\omega$ and $e\sin\omega$. So we can solve for e, p and ω, in general, after which we can get a from $p = a(1-e^2)$, including the infinite value associated with non-rectilinear parabolas, except in the indeterminate case ($p = 0$ and $e = 1$) corresponding to a rectilinear orbit.

It is obvious that the algorithm must fail in the vicinity of a rectilinear orbit. In the limit, indeed, it is only the time data that give the problem any dynamical structure. (If the indeterminacy of an orbital plane is not considered important[14], Lambert's two-vector problem can still be solved for rectilinear orbits[3].) The algorithm also fails if the determinant of the simultaneous equations is zero, in which case (following the standard theory of linear algebra) the equations are either indeterminate (infinitely many solutions) or inconsistent (no solution): when two of the u_j differ by a multiple of 2π, for example, the former conclusion follows if the corresponding r_j are equal and the latter conclusion if they are not; more generally, it is easily seen that the zero determinant arises when the three points (r_j,u_j) are collinear and e is infinite. Finally, solution of the equations (A-1) may lead to a negative value of p; this is unacceptable as it implies a repulsive force ($\mu < 0$), the orbit being the dynamically unfamiliar branch of a hyperbola.

Appendix B

DETAILS OF THE CALCPS ALGORITHM

The function of the Fortran subroutine CALCPS is to derive the components of the calculated vector $\underset{\sim}{\rho} = \underset{\sim}{c}$ at time t_2, given the vectors $\underset{\sim}{\rho}$ at times t_1 and t_3. Each vector is 'topocentric', being relative to an observer O_j (j = 1,2 or 3). The components $\underset{\sim}{R}_j$, vectors which locate the observers relative to the force centre, are held in a common block, /GIVEN/, which also specifies the sight-line unit vectors $\underset{\sim}{\lambda}_1$ and $\underset{\sim}{\lambda}_3$, the time intervals t_{12} ($= t_2 - t_1$) and t_{13} ($= t_3 - t_1$), and the gravitational constant μ (/GIVEN/ also specifies an observed sight-line $\underset{\sim}{\lambda}_2$, but this is not used by the subroutine). The subroutine has eight dummy arguments, of which the first four supply input whilst the other four define the returned output. The input arguments are: k, an assumed count of included half-revolutions, required in specifying the operation of VALAMB, the Lambert-solving subroutine; ρ_1, the assumed 'range' at time t_1 (so that $\underset{\sim}{\rho}_1$ is available as $\rho_1 \underset{\sim}{\lambda}_1$); ρ_3, similarly; and an indicator, which has to be zero if $k = 0$ or 1 but can be either zero or unity if $k \geq 2$, thereby indicating which of the two Lambert solutions (when there are two) is to be used. The first output argument specifies the number of solutions found by VALAMB; the remaining arguments are the components of the required vector $\underset{\sim}{c}$. Some subsidiary quantities, obtained during the computation, are preserved in the comon block /USEFUL/, viz α, q, i, r_1, r_2 and r_3.

The subroutine starts by obtaining $\underset{\sim}{r}$ ($= \underset{\sim}{R} + \underset{\sim}{\rho}$) at times t_1 and t_3, and hence (using both scalar product and vector product, to avoid loss of accuracy) the angle θ between $\underset{\sim}{r}_1$ and $\underset{\sim}{r}_3$; θ is initially in the interval $(0,\pi)$, but is converted to the interval $(k\pi,\overline{k+1}\pi)$ by use of the input argument k, after a preliminary conversion to $2\pi - \theta$ if k is odd. The data for VALAMB now being available, we obtain[3] radial and transverse components for the velocity at t_1. (VALAMB differs from VLAMB, the subroutine listed in Ref 3, in only one respect: the energy-equivalent quantity α, equal to μ/a where a is the semi-major axis, has become an output argument of VALAMB though only a local variable in VLAMB; this value of α is potentially more accurate, for near-parabolic orbits, than the value subsequently derived by subroutine PV3ELS.)

The subroutine next obtains the full velocity vector, $\dot{\underset{\sim}{r}}_1$, at t_1, the direction of the transverse component being supplied via the vector $(\underset{\sim}{r}_1 \wedge \underset{\sim}{r}_3) \wedge \underset{\sim}{r}_1$; when k is odd the direction of this vector must be reversed. From $\underset{\sim}{r}_1$ and $\dot{\underset{\sim}{r}}_1$ we now get the universal Keplerian elements via[14] subroutine PV3ELS, and hence propagate $\underset{\sim}{r}$ to t_2 by adding t_{12} to the element τ and calling subroutine ELS3PV; it is in this propagation that the α from VALAMB is used in preference to the α from PV3ELS.

It only remains for $\underset{\sim}{c}$ to be derived as $\underset{\sim}{r}_2 - \underset{\sim}{R}_2$.

Appendix C

LISTING OF SUBROUTINES CALCPS AND OBS3LS

C.1 Listing of subroutine CALCPS

```
      SUBROUTINE CALCPS (NHREV, RO1, RO3, IND, NUM, CX, CY, CZ)
C         (CALCulate Position along Sight-line, to
C          compare with externally given line of sight)
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      COMMON /GIVEN/ OX1, OY1, OZ1, OX2, OY2, OZ2, OX3, OY3, OZ3,
     A  EL1, EM1, EN1, EL2, EM2, EN2, EL3, EM3, EN3,  T12, T13,  GM
      COMMON /USEFUL/ AL, Q, EI, R1, R2, R3
      PARAMETER (PI=3.141592653589793D0)
      RNORM(X,Y,Z) = SQRT(X*X + Y*Y + Z*Z)
      X1 = OX1 + RO1*EL1
      Y1 = OY1 + RO1*EM1
      Z1 = OZ1 + RO1*EN1
      R1 = RNORM(X1, Y1, Z1)
      X3 = OX3 + RO3*EL3
      Y3 = OY3 + RO3*EM3
      Z3 = OZ3 + RO3*EN3
      R3 = RNORM(X3, Y3, Z3)
      CALL VECMUL (X1, Y1, Z1, X3, Y3, Z3, X13, Y13, Z13)
      TH = ATAN2(RNORM(X13, Y13, Z13), X1*X3 + Y1*Y3 + Z1*Z3)
C         (Fails only if either R1 or R2 is zero)
      M = MOD(NHREV,2)
      IF (M.EQ.1) TH = PI - TH
      TH = TH + NHREV*PI
      CALL VALAMB (GM, R1, R3, TH, T13, NUM, VR1, VT1, VR3, VT3, ALV,
     A WR1, WT1, WR3, WT3, ALW)
      IF (NUM.GT.0) THEN
        IF (IND.EQ.1) THEN
          VR1 = WR1
          VT1 = WT1
          ALV = ALW
        END IF
        CALL VECMUL (X1, Y1, Z1, X3, Y3, Z3, XN, YN, ZN)
        CALL VECMUL (XN, YN, ZN, X1, Y1, Z1, XT, YT, ZT)
        RT = RNORM(XT, YT, ZT)
        IF (M.EQ.1) RT = -RT
        IF (RT.NE.0D0) RT = 1D0/RT
        V1X = VR1*X1/R1 + VT1*XT*RT
        V1Y = VR1*Y1/R1 + VT1*YT*RT
        V1Z = VR1*Z1/R1 + VT1*ZT*RT
        CALL PV3ELS (GM, X1, Y1, Z1, V1X, V1Y, V1Z, AL, Q, EI, BOM, OM,
     A   TAU)
        CALL ELS3PV (GM, ALV, Q, EI, BOM, OM, TAU + T12, X2, Y2, Z2,
     A   W1, W2, W3)
C         (ALV gives better accuracy than AL for near-parabolic orbits)
        R2 = RNORM(X2, Y2, Z2)
C         (Only for the convergence criterion in the calling routine)
        CX = X2 - OX2
        CY = Y2 - OY2
        CZ = Z2 - OZ2
      END IF
      RETURN
      END
```

C.2 Listing of subroutine OBS3LS

```
      SUBROUTINE OBS3LS (NHREV, IND, RO1, RO3)
C        Orbit got from Observed Three Lines of Sight (angles only)
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      LOGICAL LPR, LSPEC
      DIMENSION NLDF(5), W(9), W1(9), W3(9), UW(9), RO1SQ(9), RO3SQ(9),
     A  RO1KNO(9), RO3KNO(9), RO1QU(9), RO3QU(9), RO13SQ(9), RO13QU(9)
      COMMON /GIVEN/ OX1, OY1, OZ1, OX2, OY2, OZ2, OX3, OY3, OZ3,
     A  EL1, EM1, EN1, EL2, EM2, EN2, EL3, EM3, EN3,  T12, T13,  GM
      COMMON /USEFUL/ AL, Q, EI, R1, R2, R3
      COMMON /OTHER/ LPR, MAXIT, HN, CR, NFAIL, ITNUM, NMOD, CRIT, DELNM
      COMMON /KNWNSL/ IKN, RO1KN(9), RO3KN(9), R1KNSQ(9), R3KNSQ(9)
      PARAMETER (PI=3.141592653589793D0, RADIAN=180D0/PI, ARBF=1D-3,
     A CRIVAL=1D-24)
      RNORM(X,Y,Z) = SQRT(X*X + Y*Y + Z*Z)
C
      FCOLD = 0D0
      NFAIL = 0
C
    1 NMOD = 0
      DO 14 ITNUM=0,MAXIT
      ARBFC = ARBF
      NPDF = 0
    2 NMODFY = 0
C          (NMOD for overall mods count; NMODFY is count for current
C                   iteration, limited to 2)
      ONE = 1D0
      IF (ITNUM.EQ.21) HN = 0.5D0
C          (To switch to Halley if mod NR initially)
      NFLMOD = 0
    3 CALL CALCPS (NHREV, RO1, RO3, IND, NUM, CX, CY, CZ)
      IF (ITNUM.EQ.0) THEN
        IF (RO1.EQ.RO3) THEN
          CR = EL2*CX + EM2*CY + EN2*CZ
          IF (CR.LT.0D0) THEN
            DR = ARBFC*(R1 + R3)
            CALL CALCPS (NHREV, RO1+DR, RO3+DR, IND, NUM, CX, CY, CZ)
            DELR = DR*CR/(CR - EL2*CX - EM2*CY - EN2*CZ)
            RO1 = RO1 + DELR
            RO3 = RO3 + DELR
            WRITE (2,4) RO1
    4       FORMAT ('Prelim starter replacement (both) by ', G23.15)
          END IF
        END IF
        GO TO 14
      END IF
      R10 = R1
      R30 = R3
      IF (NUM.EQ.0) THEN
        IF (ITNUM.GT.1 .AND. NFLMOD.LT.3) THEN
          NFLMOD = NFLMOD + 1
          D3 = D3/3D0
          D1 = D1/3D0
          RO1 = RO1OLD + D3
          RO3 = RO3OLD + D1
          IF (LPR) WRITE (2,5) NFLMOD, RO1, RO3
    5     FORMAT ('Lambert fail', I2, ' so cut by 2/3 to', 2E18.10)
          GO TO 3
        ELSE
```

```
          NFAIL = NFAIL + 1
          IF (NFAIL.EQ.1) THEN
C          Try to avoid Lambert-fail, by common-perpendicular starters
           X13 = OX3 - OX1
           Y13 = OY3 - OY1
           Z13 = OZ3 - OZ1
           D1 = X13*EL1 + Y13*EM1 + Z13*EN1
           D3 = X13*EL3 + Y13*EM3 + Z13*EN3
           D2 = EL1*EL3 + EM1*EM3 + EN1*EN3
           D4 = 1D0 - D2**2
           RO1 = MAX((D1 - D3*D2)/D4, 0D0)
           RO3 = MAX((D1*D2 - D3)/D4, 0D0)
           WRITE (2,6) ITNUM, RO1, RO3
    6      FORMAT ( 'Revised Starters after Lambert fail (Iteration',
     A      I3, ')'/ 'viz (common perpendicur)', E21.13, ' & ', E21.13)
           GO TO 1
          ELSE IF (NFAIL.EQ.2) THEN
C          Last chance via zero starters
           RO1 = 0
           RO3 = 0
           WRITE (2,7) ITNUM, RO1, RO3
    7      FORMAT ( 'Revised Starters after Lambert fail (Iteration',
     A      I3, ')'/ 'viz (both ZERO!)', E21.13, ' & ', E21.13)
           GO TO 1
          ELSE
           RETURN
         END IF
        END IF
       END IF
       CR = EL2*CX + EM2*CY + EN2*CZ
C         (CR  used only non-critically in convergence test)
       IF (AL.NE.0D0) A = GM/AL
       E = 1D0 - Q*AL/GM
       EIDEG = EI*RADIAN
C          (in case these  a ,  e ,  i  are to be output)
       CALL VECMUL (EL2, EM2, EN2, CX, CY, CZ, ENX, ENY, ENZ)
       CALL VECMUL (ENX, ENY, ENZ, EL2, EM2, EN2, PX, PY, PZ)
       PR = RNORM(PX, PY, PZ)
       CALL VECMUL (EL2, EM2, EN2, PX, PY, PZ, ENX, ENY, ENZ)
       ENR = RNORM(ENX, ENY, ENZ)
       IF (ENR.EQ.0D0) THEN
         CRIT = 0D0
         DELNM = 0D0
         GO TO 13
       END IF
       F = (PX*CX + PY*CY + PZ*CZ)/PR
C          (F is the P-coord; G, the N-coord, should be zero)
       FC = F
       DO 8 I=1,IKN
       RO1KNO(I) = RO1 - RO1KN(I)
       RO3KNO(I) = RO3 - RO3KN(I)
       RO1SQ(I) = RO1KNO(I)**2
       RO3SQ(I) = RO3KNO(I)**2
       RO1QU(I) = RO1SQ(I) + R1KNSQ(I)
       RO3QU(I) = RO3SQ(I) + R3KNSQ(I)
       RO13SQ(I) = RO1SQ(I) + RO3SQ(I)
       RO13QU(I) = RO1QU(I) + RO3QU(I)
    8  FC = FC*(RO13QU(I)/RO13SQ(I))**(ONE/2D0)
       LSPEC = ITNUM.GT.1
C         (To avoid Amstrad fault when ITNUM.GT.1 direct in next line)
```

```
      IF (LSPEC .AND. FC.GT.2D0*FCOLD .AND. NMODFY.LT.2) THEN
        FSUM = FC + FCOLD
        RO1 = (FC*RO1OLD + FCOLD*RO1)/FSUM
        RO3 = (FC*RO3OLD + FCOLD*RO3)/FSUM
        IF (LPR) WRITE (2,9) ITNUM, FC, FCOLD
    9   FORMAT ( 'Modify it''n', I4, '; new & old FC are', 2G18.6)
        NMODFY = NMODFY + 1
        NMOD = NMOD + 1
        GO TO 3
      END IF
C       Now get partials for (in principle) 2D-Halley
      DRO1 = ARBFC*R10
      DRO3 = ARBFC*R30
      D2RO1 = 2D0*DRO1
      D2RO3 = 2D0*DRO3
      DRO1SQ = DRO1**2
      DRO3SQ = DRO3**2
      CALL CALCPS (NHREV, RO1 - DRO1, RO3, IND, NLDF(1), CX, CY, CZ)
      FM1 = (PX*CX + PY*CY + PZ*CZ)/PR - F
      GM1 = (ENX*CX + ENY*CY + ENZ*CZ)/ENR
      CALL CALCPS (NHREV, RO1 + DRO1, RO3, IND, NLDF(2), CX, CY, CZ)
      FP1 = (PX*CX + PY*CY + PZ*CZ)/PR - F
      GP1 = (ENX*CX + ENY*CY + ENZ*CZ)/ENR
      FD1 = (FP1 - FM1)/D2RO1
      FDD1 = (FP1 + FM1)/DRO1SQ
      GD1 = (GP1 - GM1)/D2RO1
      GDD1 = (GP1 + GM1)/DRO1SQ
      CALL CALCPS (NHREV, RO1, RO3 - DRO3, IND, NLDF(3), CX, CY, CZ)
      FM3 = (PX*CX + PY*CY + PZ*CZ)/PR - F
      GM3 = (ENX*CX + ENY*CY + ENZ*CZ)/ENR
      CALL CALCPS (NHREV, RO1, RO3 + DRO3, IND, NLDF(4), CX, CY, CZ)
      FP3 = (PX*CX + PY*CY + PZ*CZ)/PR - F
      GP3 = (ENX*CX + ENY*CY + ENZ*CZ)/ENR
      FD3 = (FP3 - FM3)/D2RO3
      FDD3 = (FP3 + FM3)/DRO3SQ
      GD3 = (GP3 - GM3)/D2RO3
      GDD3 = (GP3 + GM3)/DRO3SQ
      CALL CALCPS (NHREV, RO1 + DRO1, RO3 + DRO3, IND, NLDF(5),
     A  CX, CY, CZ)
      F13 = (PX*CX + PY*CY + PZ*CZ)/PR - F
      G13 = (ENX*CX + ENY*CY + ENZ*CZ)/ENR
      ROFAC = DRO1/DRO3
      FD13 = F13/(DRO1*DRO3) - (FD1/DRO3 + FD3/DRO1)
     A - 0.5D0*(FDD1*ROFAC + FDD3/ROFAC)
      GD13 = G13/(DRO1*DRO3) - (GD1/DRO3 + GD3/DRO1)
     A - 0.5D0*(GDD1*ROFAC + GDD3/ROFAC)
      DO 11 I=1,5
      IF (NLDF(I).EQ.0) THEN
        NPDF = NPDF + 1
        IF (NPDF.LE.3) THEN
          ARBFC = ARBFC/10D0
          IF (LPR) WRITE (2,10)
   10     FORMAT ('Reduce ARBFC')
          GO TO 2
         ELSE
          NFAIL = 1
          RETURN
        END IF
      END IF
   11 CONTINUE
```

```
      DO 12 I=1,IKN
      W(I) = 1D0/RO13SQ(I) - 1D0/RO13QU(I)
      UW(I) = ONE*W(I) - (2D0/RO13SQ(I) + 2D0/RO13QU(I))
      W1(I) = W(I)*RO1KNO(I)
      W3(I) = W(I)*RO3KNO(I)
      FD1 = FD1 - ONE*F*W1(I)
      FD3 = FD3 - ONE*F*W3(I)
      FDD1 = FDD1 - ONE*(2D0*FD1*W1(I) + W(I)*F*(1D0 + RO1SQ(I)*UW(I)))
      GDD1 = GDD1 - ONE*2D0*W1(I)*GD1
      FDD3 = FDD3 - ONE*(2D0*FD3*W3(I) + W(I)*F*(1D0 + RO3SQ(I)*UW(I)))
      GDD3 = GDD3 - ONE*2D0*W3(I)*GD3
      FD13 = FD13 - ONE*(FD3*W1(I) + FD1*W3(I) + W(I)*F*
     A RO1KNO(I)*RO3KNO(I)*UW(I))
   12 GD13 = GD13 - ONE*(W1(I)*GD3 + W3(I)*GD1)
      DEL = FD1*GD3 - FD3*GD1
      D3NR = - GD3*F/DEL
      D1NR = GD1*F/DEL
      FD1H = FD1 + HN*(FDD1*D3NR + FD13*D1NR)
      FD3H = FD3 + HN*(FD13*D3NR + FDD3*D1NR)
      GD1H = GD1 + HN*(GDD1*D3NR + GD13*D1NR)
      GD3H = GD3 + HN*(GD13*D3NR + GDD3*D1NR)
      DELH = FD1H*GD3H - FD3H*GD1H
      D3 = -GD3H*F/DELH
      D1 = GD1H*F/DELH
      DELNM = DEL/(DABS(FD1*GD3) + DABS(FD3*GD1))
      IF (ABS(DELNM).LT.1D-3) THEN
        D3 = SIGN(SQRT(ABS(D3NR*D3)), D3NR)
        D1 = SIGN(SQRT(ABS(D1NR*D1)), D1NR)
      END IF
      RO1OLD = RO1
      RO3OLD = RO3
      RO1 = RO1 + D3
      RO3 = RO3 + D1
      DELNM = DELH/(DABS(FD1H*GD3H) + DABS(FD3H*GD1H))
      DEN = DMAX1(CR, R2)
      IF (A.GT.0D0) DEN = DMAX1(DEN, T12*SQRT(ABS(AL*(2D0*A-R2))/R2))
      CRIT = F/DEN
   13 IF (CRIT**2.LT.CRIVAL) GO TO 15
      FCOLD = FC
   14 CONTINUE
      NFAIL = -1
      RETURN
   15 NFAIL = 0
      AL = A
      Q = E
      EI = EIDEG
      RETURN
      END
```

Table 1

DATA FOR EXAMPLE OF HERRICK

Times	0.325593 (t_{12}) and	0.701944 (t_{13})	
$\underset{\sim}{R}_1$	0.7000687	0.6429399	0.2789211
$\underset{\sim}{R}_2$	0.4306907	0.8143496	0.3532745
$\underset{\sim}{R}_3$	0.0628371	0.9007098	0.3907417
$\underset{\sim}{\lambda}_1$	0.9028975	0.0606048	0.4255621
$\underset{\sim}{\lambda}_2$	0.9224764	0.0518570	0.3825549
$\underset{\sim}{\lambda}_3$	0.9347684	0.0802269	0.3460802

(Light-corrected times not used but are 0.325578 and 0.701903)

Table 2

RESULTS FOR HERRICK EXAMPLE

Sol	N	ρ_1	ρ_2	ρ_3	ε	e
1	4	2.399197226489	2.563703947213	2.824544883197	3.1E-15	0.049
2	5	-0.7972181001259	-0.9956194556367	-1.0812024691879	0.7E-15	0.858
3	10	-0.0003632101136	-0.0001443130763	0.0001663085092	0.5E-15	0.015

Table 3

RESULTS FOR HERRICK-BASED EXAMPLE WITH LINEAR DEPENDENCE

Sol	ρ_1	ρ_2	ρ_3	e
1	2.39919799711	2.59571077780	2.82454554661	0.05
2	-3.45697670528	-3.68976500360	-3.94562359213	4.57

Table 4

RESULTS FOR HERRICK-BASED EXAMPLE WITH RANK EQUALS 1

Sol	ρ_1	ρ_2	ρ_3	e
1 and 2	±1.299928870	±1.334032020	±1.368260609	0.05
3 and 4	±3.790506796	±2.223428041	∓3.902205940	355

AGENCY

Overall security classification of this page

Table 5

RESULTS FOR HERRICK-BASED EXAMPLE WITH NEIGHBOURING SOLUTIONS

Sol	ρ_1	ρ_2	ρ_3	e
1	2.3991972	2.5633570	2.8245448	0.05
2	2.3990000	2.5630000	2.8240000	0.05
3	-0.677952469654	-3.106437154553	-5.812102316078	40

Table 6

DATA FOR EXAMPLE OF EXTREME SEVERITY

Times	1.570796327 (t_{12})	and	3.141592654 (t_{13})
$\underset{\sim}{R}_1$	0.0	-1.0	0.0
$\underset{\sim}{R}_2$	1.0	0.0	0.0
$\underset{\sim}{R}_3$	0.0	1.0	0.0
$\underset{\sim}{\lambda}_1$	0.0	1.0	0.0
$\underset{\sim}{\lambda}_2$	0.0	0.0	1.0
$\underset{\sim}{\lambda}_3$	0.0	-1.0	0.0

Table 7

PARTICULAR RESULTS (WITH $\rho_2 = 0$) FOR EXTREME EXAMPLE

Sol	k	ρ_1	ρ_3	e	Type
1	0	0.00000000000	0.00000000000	0	S D
2	0	2.00000000000	2.00000000000	0	S R
3	0	0.56960576485	1.43039423515	0.617	U D
4	0	1.43039423515	0.56960576485	0.617	U R
5	1	1.43039423515	0.56960576485	0.617	U D
6	1	0.56960576485	1.43039423515	0.617	U R
7	1	1.309904240665	1.309904240665	0.690	S D
8	1	0.690095759335	0.690095759335	0.690	S R

Code: S – symmetric; U – unsymmetric; D – direct; R – retrograde
(k is based on t_{12} not t_{13})

Table 8

DATA FOR EXAMPLE INVOLVING RECTILINEAR ORBIT

Times	1.047197552 (t_{12})	and	2.960420507 (t_{13})
$\underset{\sim}{R}_1$	1.0	0.0	0.0
$\underset{\sim}{R}_2$	0.0	1.0	0.0
$\underset{\sim}{R}_3$	0.0	0.0	0.0
$\underset{\sim}{\lambda}_1$	-1.0	0.0	0.5
$\underset{\sim}{\lambda}_2$	0.0	-1.0	1.5
$\underset{\sim}{\lambda}_3$	0.0	0.0	2.0

Table 9

RESULTS FOR RECTILINEAR-ORBIT EXAMPLE

Sol	ρ_2	ρ_3	a	e
1	1.8027756377	1.9999999978	0.999999999	1.000
2	1.8027756377	3.2004916441	14.683154273	1.000
3	1.3950850319	0.5000000000	0.605502151	0.984
4	1.2038678317	0.5000000000	0.605502151	0.979
5	0.2904162720	0.5000000000	0.605502151	0.743
6	0.2253738553	0.5000000000	0.605502151	0.481
7	0.0468631594	-0.7176281374	0.731442384	0.441
8	0.3982547067	-0.1703288931	0.559537113	0.739

ρ_1 = 1.1180339887 for every solution

Table 10

DATA FOR FIRST EXAMPLE OF ESCOBAL

Times (s)	3296.449 (t_{12})	and	3450.505 (t_{13})
$\underset{\sim}{R}_1$	1094.6803	5355.5873	-3275.6560
$\underset{\sim}{R}_2$	-4718.4298	-2606.2925	3401.1164
$\underset{\sim}{R}_3$	-4688.8537	-2659.1335	3401.1164
$\underset{\sim}{\lambda}_1$	-0.12989024	0.92763536	-0.35017305
$\underset{\sim}{\lambda}_2$	0.23466879	-0.85577972	0.46105491
$\underset{\sim}{\lambda}_3$	0.42481281	-0.75210323	0.50385991

Table 11

RESULTS FOR FIRST ESCOBAL EXAMPLE

Sol	k	ρ_1	ρ_2	ρ_3	e
1	0	-7833.0603681592	2228.5012290292	1257.6207462956	0.79
2	0	-12478.3464866953	2252.4162998993	1213.7982019823	0.77
3	1	-14142.9610065184	-2990.0836578326	-1830.0797713851	0.28
4	1	-3283.1722962962	834.0149392002	-361.0803280551	0.43
5	1	-3202.2860993118	3718.7361430846	3891.0612321812	0.47

Table 12

REVISED DATA FOR FIRST EXAMPLE OF ESCOBAL

Times	0.0381533 (t_{12}) and 0.0399364 (t_{13})		
$\underset{\sim}{R}_1$	0.16606957	0.84119785	-0.51291356
$\underset{\sim}{R}_2$	-0.73815134	-0.41528280	0.53035336
$\underset{\sim}{R}_3$	-0.73343987	-0.42352540	0.53037164
$\underset{\sim}{\lambda}_1$	-0.92475472	-0.37382824	-0.07128226
$\underset{\sim}{\lambda}_2$	0.80904274	-0.55953385	0.17992142
$\underset{\sim}{\lambda}_3$	0.85044131	-0.49106628	0.18868888

Table 13

RESULTS FOR EXAMPLE WITH REVISED DATA

Sol	k	ρ_1	ρ_2	ρ_3	e
1	0	7.508030354109	2.970622569286	3.290782845481	10.09
2	0	3.591011270710	1.883891729169	2.031860787010	1.45
3	0	1.100072662216	1.518998485736	1.578206979185	0.18
4	0	0.014616956287	-0.446474185383	-0.291261277917	0.45
5	1	-0.513529550563	1.194401144814	1.313589989963	0.27
6	1	2.480762217620	-0.843311721142	-0.970019178693	0.90
7	1	0.225016498047	0.439481228548	0.235086077861	0.13

Table 14

DATA FOR SECOND EXAMPLE OF ESCOBAL

Times (s)	60.459 (t_{12}) and 128.516 (t_{13})		
$\underset{\sim}{R}_1$	3941.7283	-3291.9961	3769.7326
$\underset{\sim}{R}_2$	3956.2035	-3274.5861	3769.7326
$\underset{\sim}{R}_3$	3972.4058	-3254.9120	3769.7326
$\underset{\sim}{\lambda}_1$	-0.25745288	-0.93936322	0.22652759
$\underset{\sim}{\lambda}_2$	-0.15323536	-0.96109718	0.22980675
$\underset{\sim}{\lambda}_3$	-0.00248708	-0.97435020	0.22502333

Table 15

SOLUTION FOR SECOND ESCOBAL EXAMPLE

ρ_1	ρ_2	ρ_3	e
2728.446508110	2404.741709	2061.116114538	0.197

Table 16

DATA FOR EXAMPLE OF LANE

Times (s)	2.2 (t_{12}) and 4.4 (t_{13})		
$\underset{\sim}{R}_1$	0.51338024	-0.78261472	0.35225232
$\underset{\sim}{R}_2$	0.89263524	0.28086002	0.35277012
$\underset{\sim}{R}_3$	-0.02703285	0.93585748	0.35152067
$\underset{\sim}{\lambda}_1$	0.76526944	0.12314580	0.63182102
$\underset{\sim}{\lambda}_2$	-0.21266402	-0.54295751	0.81238609
$\underset{\sim}{\lambda}_3$	0.52029946	-0.39083440	0.75930030

Table 17

RESULTS FOR LANE EXAMPLE

Sol	(*)	k	ρ_1	ρ_2	ρ_3	a	e
1	*	0	0.837595099	-34.922279846	-6.962795558	18.113	0.932
	*	0	-1.201563236	35.251656399	0.665839729	17.823	0.976
2	*	1	6.226195556	-30.966211020	5.641736738	17.915	0.735
3		1	3.655412566	34.624802268	7.925507622	17.984	0.925
4	*	2	6.224498976	-30.436814747	5.642945981	17.651	0.731
	*	2	-4.592244879	-33.156970920	2.297947486	17.575	0.891
5	*	2	7.043008437	-0.073841273	7.177034987	11.543	0.990
6	*	2	3.618416215	1.349813283	5.641096571	11.376	0.996
	*	2	-14.976761959	0.483717653	-14.397094839	12.367	0.913
	*	3	-1.201387949	35.093558679	0.665728275	17.744	0.975
7		3	0.357582617	0.756876743	1.101195403	11.221	0.893
8		4	0.357552418	0.756881606	1.101182380	11.184	0.893
	*	4	0.715681869	-16.074598191	-5.428181352	8.674	0.866
	*	4	-1.160059994	16.725808742	0.639626121	8.568	0.949
9	*	5	7.022114061	-0.081679948	7.160031641	10.855	0.989
10	*	5	-13.041724491	0.485214899	-12.461028211	10.174	0.894
11	*	5	3.623893984	1.352910753	5.444619148	11.028	0.996
12	*	5	6.113063159	-12.183403755	5.735762464	8.619	0.457
13	12	5	3.814908783	15.725170236	7.512753553	8.645	0.838
	*	5	-1.347402915	-16.301985468	0.761755147	8.571	0.912
14	*	6	6.109805387	-11.896567152	5.739037781	8.479	0.449
	10	6	3.821959154	15.328237166	7.495758721	8.453	0.835
	*	6	-1.348561987	-16.215835054	0.762686952	8.528	0.912
	*	6	-3.724080611	-15.123455751	2.046411908	8.469	0.797
15	*	6	6.840279880	-0.161036698	7.010392376	7.283	0.979
16	*	6	-10.335455013	0.490354274	-9.752859758	7.523	0.856
	*	6	3.727367055	1.418037228	5.513414148	7.176	0.992
17	*	7	-1.159714114	16.650076229	0.639409390	8.530	0.949
18	*	7	0.709255856	-15.576923099	-5.359701407	8.425	0.863
19		7	0.351984180	0.757780461	1.098766810	7.069	0.831
20		8	0.351931374	0.757789006	1.098743765	7.045	0.830
21	*	8	-1.128900254	11.766048510	0.620247182	6.095	0.928
22	*	8	0.630536717	-11.059572708	-4.600174082	6.158	0.818
23	*	9	6.800767940	-0.181487682	6.977376208	6.826	0.976
	*	9	3.738405100	1.425825307	5.521031316	6.937	0.991
	*	9	-9.320041449	0.492096117	-8.736536971	6.634	0.837
24	*	9	6.041019575	-7.029150141	5.824139576	6.138	0.319
	11	9	3.941520416	10.495594171	7.219703948	6.151	0.777
	*	9	-3.167853059	-10.595528041	1.839139953	6.129	0.746
25	*	10	6.037018187	-6.790258941	5.830613202	6.025	0.316
	9	10	3.951865064	10.192743011	7.196709213	6.012	0.773
	*	10	-1.463579232	-11.220014314	0.853605966	6.065	0.866
	*	10	-3.137129497	-10.416224254	1.826400002	6.035	0.744
26	*	10	6.649249752	-0.274636956	6.848377683	5.568	0.963
27	*	10	-8.122468268	0.494228051	-7.537698084	5.650	0.808
	*	10	3.833093073	1.501066889	5.589037806	5.484	0.986
28	*	11	-1.128430803	11.712192739	0.619957706	6.068	0.928
29		11	0.346938814	0.758598749	1.096553397	5.395	0.779
30		12	0.346863530	0.758610987	1.096520193	5.377	0.778
31	*	12	-1.102521977	9.304018293	0.604116918	4.870	0.910
	*	12	0.561286390	-8.584433580	-4.035301284	4.914	0.779

Table 17 (continued)

Sol	(*)	k	ρ_1	ρ_2	ρ_3	a	e
32	*	13	6.585485324	-0.322853386	6.792610550	5.196	0.956
	*	13	-7.470360864	0.495379975	-6.884792007	5.142	0.789
	*	13	3.851548188	1.517783836	5.602923783	5.285	0.985
33	*	13	5.997706056	-4.347903871	5.924152801	4.912	0.379
34	*	13	-2.655894748	-8.293223274	1.606338565	4.890	0.720
	8	13	4.054258362	7.743637603	6.973355905	4.918	0.742
	*	13	-1.623200305	-8.748732237	0.974365187	4.875	0.818
35	*	14	5.994774086	-4.104469948	5.937627103	4.806	0.397
	6	14	4.068199886	7.467561286	6.943186244	4.801	0.739
36	*	14	6.451436991	-0.450122661	6.671153554	4.607	0.936
37	*	14	3.942942189	1.613639602	5.675548297	4.535	0.979
38	*	14	-6.777869246	0.496535104	-6.191348804	4.621	0.765
39	*	15	-1.101949584	9.260918768	0.603770189	4.849	0.909
40		15	0.342136019	0.759381161	1.094424627	4.454	0.732
41		16	0.342037425	0.759397259	1.094380704	4.438	0.731
42	*	16	-1.079108596	7.789602380	0.590064393	4.119	0.893
43	*	16	0.501483394	-7.070867044	-3.610550759	4.152	0.745
44	*	17	6.343610751	-0.591457361	6.567314506	4.272	0.912
	*	17	3.972396098	1.650287279	5.700608287	4.354	0.976
	*	17	-6.313364048	0.497225404	-5.726136069	4.282	0.747
45	*	17	5.997658907	-2.494991767	6.067616849	4.167	0.579
	5	17	4.158488025	5.876436995	6.743793054	4.163	0.740
46	*	18	6.011121742	-2.122298719	6.115187076	4.047	0.636
	2	18	4.176734108	5.581862791	6.701400977	4.055	0.744
47	*	18	6.207705855	-0.965299003	6.422515298	3.991	0.863
48	*	18	4.063081080	1.788939357	5.784918395	3.917	0.966
	*	18	-5.856182158	0.497788099	-5.268195958	3.958	0.726
49	*	19	-1.078445124	7.753045732	0.589670316	4.101	0.893
50	*	19	0.493797645	-6.906032263	-3.559675134	4.069	0.741
		19	0.337447584	0.760148272	1.092326074	3.838	0.689
51		20	0.337324382	0.760168475	1.092270656	3.825	0.688
52	*	20	-1.057767132	6.746716271	0.577516105	3.603	0.877
	*	20	0.448114374	-6.034823566	-3.273000619	3.629	0.715
53	*	21	4.112046684	1.887395487	5.836799521	3.740	0.958
	*	21	-5.503601434	0.498101920	-4.914984311	3.712	0.708
	*	22	4.274920242	3.911319258	6.417826465	3.537	0.805
54	*	22	4.205884803	2.168938828	5.960356096	3.484	0.936
55	*	23	-1.057020072	6.714608302	0.577081883	3.587	0.877
56	*	23	0.440318364	-5.902503717	-3.226624705	3.562	0.710
57		23	0.332803487	0.760911435	1.090227447	3.399	0.649
58		24	0.332654056	0.760936046	1.090159591	3.387	0.648
59	*	24	0.399485904	-5.274048002	-2.994890332	3.244	0.688
60	*	24	-1.037971837	5.976212100	0.566137052	3.224	0.863
	*	26	-4.649742490	0.498107348	-4.059413059	3.138	0.654
61	*	27	-1.037146045	5.947344048	0.565668353	3.209	0.862
62		27	0.328158919	0.761678034	1.088108797	3.067	0.612
63		28	0.327981359	0.761707408	1.088027409	3.056	0.610
64	*	28	-1.019378287	5.378830225	0.555710684	2.930	0.848
65	*	31	-1.018476840	5.352434071	0.555212231	2.917	0.848
66		31	0.323481873	0.762453454	1.085955331	2.806	0.576
67		32	0.323274021	0.762487997	1.085859163	2.796	0.575
68	*	32	-1.001743938	4.899080435	0.546087591	2.696	0.834
69	*	32	0.312547488	-4.219866927	-2.557300082	2.710	0.640
70	*	35	-1.000768600	4.874634181	0.545563510	2.684	0.834

Table 17 (continued)

Sol	(*)	k	ρ_1	ρ_2	ρ_3	σ	e
71	*	35	0.304676663	-4.139423022	-2.521034654	2.669	0.636
72		35	0.318747488	0.763241989	1.083755045	2.594	0.542
73		36	0.318506908	0.763282158	1.083642688	2.584	0.540
74	*	36	-0.984888657	4.503305085	0.537160038	2.503	0.821
75	*	39	-0.983840152	4.480435696	0.536614080	2.492	0.820
76	*	39	0.265155717	-3.765528113	-2.346489963	2.480	0.615
77		39	0.313935103	0.764047312	1.081497480	2.418	0.509
78		40	0.313659070	0.764093622	1.081367346	2.409	0.508
79	*	40	-0.968672954	4.169801453	0.528847979	2.342	0.808
80	*	43	-0.967551130	4.148231738	0.528283677	2.331	0.807
81	*	43	0.227713581	-3.451828053	-2.192013475	2.320	0.595
82		43	0.309026563	0.764872748	1.079172993	2.269	0.478
83		44	0.308712039	0.764925780	1.079023297	2.261	0.476
84	*	44	-0.952985023	3.883892287	0.521090781	2.204	0.795
	*	44	0.199843620	-3.240146383	-2.083382418	2.213	0.580
85	*	47	-0.951788951	3.863409707	0.520511577	2.194	0.794
86		47	0.304005170	0.765721454	1.076772301	2.141	0.448
87		48	0.303648769	0.765781860	1.076601034	2.133	0.446
88	*	48	-0.937732537	3.635264647	0.513842069	2.085	0.783
89	*	51	-0.936460559	3.615701948	0.513251414	2.075	0.782
90		51	0.298854976	0.766596536	1.074286150	2.030	0.419
91		52	0.298452924	0.766665048	1.074091051	2.022	0.416
92	*	52	-0.922837189	3.416441108	0.507066370	1.980	0.770
	*	52	0.132892246	-2.794509822	-1.842549332	1.988	0.545
93	*	55	-0.921486942	3.397663738	0.506467818	1.971	0.769
94		55	0.293560276	0.767501140	1.071705075	1.932	0.390
95		56	0.293108360	0.767578587	1.071483593	1.925	0.388
96	*	56	-0.908230900	3.221855968	0.500736884	1.888	0.758
97	*	59	-0.906799304	3.203753597	0.500134182	1.880	0.757
	*	59	0.093603813	-2.567250497	-1.713266557	1.873	0.525
98		59	0.288105208	0.768438530	1.069019185	1.846	0.363
99		60	0.287598722	0.768525846	1.068768439	1.838	0.361
100	*	60	-0.893853078	3.047273122	0.494833972	1.806	0.746
	*	60	0.070745719	-2.444765037	-1.641800951	1.812	0.513
101	*	63	-0.892336306	3.029753742	0.494231153	1.798	0.745
	*	63	0.063090345	-2.405198903	-1.618455547	1.792	0.509
102		63	0.282473429	0.769412155	1.066217981	1.768	0.337
103		64	0.281907098	0.769510398	1.065934703	1.761	0.335
104	*	64	-0.879648556	2.889406047	0.489344138	1.732	0.734
105	*	67	-0.878041979	2.872391752	0.488745624	1.725	0.733
	*	67	0.033515032	-2.258672995	-1.530936347	1.719	0.494
106		67	0.276647808	0.770425716	1.063290167	1.698	0.312
107		68	0.276015712	0.770536090	1.062970648	1.691	0.310
108	*	68	-0.865565978	2.745662077	0.484259370	1.666	0.723
109	*	71	-0.863864093	2.729085979	0.483670083	1.658	0.721
110		71	0.270610146	0.771483233	1.060223463	1.635	0.288
111		72	0.269905613	0.771607111	1.059863471	1.628	0.286
112	*	72	-0.851556489	2.613965778	0.479576745	1.605	0.711
	*	72	-0.015618104	-2.035109804	-1.394441154	1.610	0.469
113	*	75	-0.849752811	2.597769698	0.479002235	1.598	0.709
114		75	0.264340873	0.772589122	1.057004391	1.577	0.266
115		76	0.263556365	0.772728077	1.056599086	1.570	0.263
116	*	76	-0.837572606	2.492634106	0.475298242	1.550	0.699
	*	77	-2.045823549	0.460902913	-1.452068027	1.557	0.315

Table 17 (continued)

Sol	(*)	k	ρ_1	ρ_2	ρ_3	a	e
117	*	79	-0.835659542	2.476766821	0.474744845	1.543	0.698
118		79	0.257818732	0.773748278	1.053618040	1.524	0.244
119		80	0.256945701	0.773904122	1.053161864	1.517	0.242
120	*	80	-0.823567215	2.380286392	0.471430754	1.499	0.688
121	*	83	-0.821535896	2.364702271	0.470905773	1.492	0.686
	*	83	-0.077373558	-1.783376006	-1.237401050	1.489	0.436
122		83	0.251020425	0.774966179	1.050047795	1.475	0.224
123		84	0.250049132	0.775141006	1.049534337	1.469	0.222
124	*	84	-0.809492613	2.275778251	0.467986261	1.452	0.677
	*	84	-0.095995112	-1.712848879	-1.193004520	1.456	0.427
	*	86	-1.826146530	0.445813164	-1.234536042	1.440	0.264
125	*	87	-0.807332692	2.260436185	0.467498189	1.445	0.675
126		87	0.243920197	0.776249002	1.046275011	1.430	0.206
127		88	0.242839502	0.776445246	1.045696843	1.424	0.204
128	*	88	-0.795299564	2.178152287	0.464982180	1.409	0.665
	*	89	-1.777936005	0.441607788	-1.187077097	1.414	0.252
129	*	91	-0.792998953	2.163014835	0.464540975	1.402	0.663
130		91	0.236489353	0.777603767	1.042278621	1.388	0.191
131		92	0.235286448	0.777824277	1.041627090	1.382	0.188
132	*	92	-0.780936303	2.086600793	0.462441898	1.369	0.654
	*	94	-1.668066947	0.430398831	-1.079489903	1.358	0.225
133	*	95	-0.778480854	2.071633479	0.462059323	1.362	0.652
134		95	0.228695673	0.779038522	1.038034663	1.350	0.177
135		96	0.227355749	0.779286651	1.037299634	1.343	0.176
136	*	96	-0.766347457	2.000437161	0.460395512	1.331	0.643
	*	96	-0.172556375	-1.444150120	-1.023826325	1.334	0.384
	*	97	-1.624928274	0.425261735	-1.037520646	1.336	0.215
	*	98	-1.595301725	0.421450761	-1.008807884	1.321	0.207
137		100	0.219008516	0.780842294	1.032685210	1.307	0.166
138	*	100	-0.751472814	1.919073712	0.458880813	1.296	0.632
	*	100	-0.197609980	-1.362544221	-0.972907389	1.299	0.370
139	*	103	-0.748654848	1.904352319	0.458660235	1.290	0.630
140		103	0.211868796	0.782186734	1.028690006	1.280	0.161
141		104	0.210200187	0.782502824	1.027749901	1.273	0.160
142	*	104	-0.736245876	1.842004317	0.457944559	1.263	0.621
	*	106	-1.459787797	0.400413349	-0.878976262	1.254	0.174
143	*	107	-0.733213603	1.827362076	0.457833561	1.257	0.618
144		107	0.202746104	0.783923785	1.023520769	1.248	0.158
145		108	0.200879223	0.784281981	1.022454064	1.242	0.158
146	*	108	-0.720592124	1.768790626	0.457644128	1.233	0.610
147	*	111	-0.717317797	1.754199894	0.457667237	1.227	0.607
148		111	0.193078982	0.785788878	1.017964720	1.218	0.159
149		112	0.190985433	0.786196191	1.016750927	1.212	0.160
150	*	112	-0.704426850	1.699051034	0.458049635	1.204	0.599
	*	113	-1.360037030	0.379949492	-0.785652246	1.207	0.151
151	*	115	-0.700877113	1.684484848	0.458236844	1.198	0.596
152	*	115	0.182802135	0.787800242	1.011970601	1.190	0.164
153	*	116	0.180447739	0.788265325	1.010584736	1.184	0.165
154	*	116	-0.687652429	1.632451704	0.459246671	1.176	0.588
155	*	119	-0.683786826	1.617883355	0.459634940	1.170	0.586
156	*	119	0.171838035	0.789980072	1.005477002	1.163	0.172
157	*	120	0.169181167	0.790513751	1.003888239	1.157	0.174
158	*	120	-0.670154754	1.568699117	0.461339872	1.151	0.577
159	*	123	-0.665923718	1.554101619	0.461975134	1.145	0.575

Table 17 (concluded)

Sol	(*)	k	ρ_1	ρ_2	ρ_3	a	e
160	*	123	0.160093413	0.792355785	0.998409421	1.138	0.183
161	*	124	0.157082674	0.792971825	0.996579207	1.132	0.186
162	*	124	-0.651798488	1.507533642	0.464457633	1.126	0.567
163	*	127	-0.647140559	1.492879012	0.465397530	1.120	0.564
164	*	127	0.147454338	0.794961820	0.990676151	1.114	0.196
165	*	128	0.144025225	0.795678041	0.988555497	1.108	0.200
166	*	128	-0.632420551	1.448723629	0.468758453	1.103	0.556
167	*	131	-0.627258390	1.433981798	0.470076152	1.097	0.554
168	*	131	0.133779141	0.797842283	0.982162363	1.092	0.211
169	*	132	0.129849099	0.798682270	0.979687813	1.086	0.215
170	*	132	-0.611820843	1.392059400	0.474439728	1.081	0.546
	*	133	-0.409083531	-0.757495186	-0.624321289	1.082	0.239
	*	134	-0.421522572	-0.725458378	-0.608318388	1.074	0.231
	*	134	-1.044133746	0.261006615	-0.519161800	1.073	0.104
171	*	135	-0.606055364	1.377196436	0.476229412	1.075	0.543
172	*	135	0.118887840	0.801054982	0.972721290	1.070	0.228
173	*	136	0.114348667	0.802050790	0.969808722	1.064	0.233
174	*	136	-0.589748502	1.337346183	0.481750361	1.060	0.536
	*	138	-0.456012568	-0.638399631	-0.566684835	1.053	0.211
175	*	139	-0.583249809	1.322321254	0.484135476	1.054	0.533
176	*	139	0.102545729	0.804677841	0.962160577	1.050	0.246
177	*	140	0.097251311	0.805874489	0.958695185	1.044	0.251
178	*	140	-0.565880464	1.284394426	0.491009803	1.040	0.526
	*	141	-0.478658352	-0.582675730	-0.541590562	1.041	0.198
	*	142	-0.496333814	-0.540003665	-0.523273963	1.033	0.188
179	*	143	-0.558473076	1.269154317	0.494156110	1.034	0.523
180	*	143	0.084436557	0.808819625	0.950220050	1.031	0.264
181	*	144	0.078181903	0.810283008	0.946039225	1.024	0.271
182	*	144	-0.539785836	1.233004419	0.502638710	1.021	0.516
	*	145	-0.524013061	-0.474702955	-0.496888534	1.022	0.175
183	*	147	-0.531222838	1.217473107	0.506776515	1.015	0.513
184	*	147	0.064115851	0.813639040	0.936533174	1.012	0.284
185	*	148	0.056598393	0.815471022	0.931394954	1.006	0.291
186	*	148	-0.510861814	1.182938727	0.517212684	1.003	0.507
	*	150	-0.672969313	-0.162159519	-0.406393714	0.996	0.128
187	*	151	-0.500775280	1.166996116	0.522678500	0.997	0.503
188	*	151	0.040922622	0.819381779	0.920554498	0.994	0.305
189	*	152	0.031663089	0.821752059	0.914073195	0.988	0.313
190	*	152	-0.478205899	1.133866609	0.535567466	0.985	0.497
191	*	155	0.013792194	0.826461040	0.901406637	0.977	0.327
192	*	156	0.001947228	0.829685243	0.892900917	0.971	0.335
193	*	159	-0.019213373	0.835664974	0.877498653	0.961	0.350

* An asterisk in column 2 indicates an impossible solution; a number in column 2 refers to a solution found by Lane.

LIST OF SYMBOLS

a	semi-major axis
$\underset{\sim}{c}$	calculated version of the vector $\underset{\sim}{\rho}_2$
c_j	coefficients (Gauss's method only)
C	centre of force
D	determinant (Jacobian) of partial derivatives
e	eccentricity
E	eccentric anomaly
f,g	target functions to be nulled ($g = 0$ at start of each iteration)
F,G	f,g multiplied by factors γ/β
i	inclination of orbit (to the xy-plane)
j	index (1-3) for the observations (often omitted for $j = 2$)
k	number of half-revolutions spanned by t_{13}
kn	suffix meaning 'known'
m,M	quantities used in Laplace's method
M	mean anomaly
n	mean motion
N	number of iterations to convergence
NR	Newton-Raphson
O	observing site
p	orbit parameter $a(1-e^2)$
P	position in orbit
q	pericentre radius
$\underset{\sim}{r}$	vector $\overrightarrow{CP}$
$\underset{\sim}{R}$	vector $\overrightarrow{CO}$
t	time
t_{jk}	$t_k - t_j$
T	orbital period (ellipses)
v	true anomaly
$\underset{\sim}{V}$	velocity vector
w	$1/\beta^2 - 1/\alpha^2$
x,y,z	components of $\underset{\sim}{r}$
x,y	alternative notation for ρ_1 and ρ_3
X,Y,Z	components of $\underset{\sim}{R}$
X,Y	known solution for ρ_1 and ρ_3
α	μ/a
β	$(\xi^2 + \eta^2)^{1/2}$
γ	$\left(\beta^2 + r_{1,\mathrm{kn}}^2 + r_{3,\mathrm{kn}}^2\right)^{1/2}$

LIST OF SYMBOLS (concluded)

γ	coefficient (Gauss's method only)
δ	increment
ε	criterion quantity for convergence
ζ,η	see ξ
θ	angle P_1CP_3
$\underline{\lambda}$	observed-direction affine vector
μ	gravitation constant for the force centre
ξ,η	$x-X$ and $y-Y$, when $x=\rho_1$ etc
ξ,η,ζ	components of $\underline{\rho}$ $(=\rho\underline{\lambda})$ or $\underline{\varrho}$
$\underline{\rho}$	vector OP
τ	t at pericentre
ϕ,ψ	angles used in Laplace's method (and Gauss's)
ω	argument of pericentre
Ω	longitude (or right ascension) of the node

REFERENCES

No.	Author	Title, etc
1	R.H. Gooding A.W. Odell	*A monograph on Kepler's equation.* RAE Technical Report 85080 (1985) Also (shortened) in *Celestial Mech.*, **38**, 307-334 (1986)
2	R.H. Gooding	*Solution of the hyperbolic Kepler's equation.* RAE Technical Report 87042 (1987) Also (shortened) in *Celestial Mech.*, **44**, 267-282 (1988)
3	R.H. Gooding	*On the solution of Lambert's orbital boundary-value problem.* RAE Technical Report 88027 (1988) Also (shortened) in *Celestial Mech.*, **48**, 145-165 (1990)
4	R.H. Gooding R.J. Tayler	*A PROP3 users' manual.* RAE Technical Report 68299 (1968)
5	R.H. Gooding	*The evolution of the PROP6 orbit determination program, and related topics.* RAE Technical Report 74164 (1974)
6	F.R. Moulton	*An introduction to Celestial Mechanics.* New York, Macmillan (1902, 1914 and 1958) and New York, Dover (1970)
7	H.C. Plummer	*An introductory treatise on Dynamical Astronomy.* New York, Dover (1918 and 1960)
8	J.M.A. Danby	*Fundamentals of Celestial Mechanics.* New York, Macmillan (1962) and Richmond (Virginia), Willmann-Bell (1988)
9	R. Deutsch	*Orbital dynamics of space vehicles.* Englewood Cliffs (New Jersey), Prentice-Hall (1963)
10	P.R. Escobal	*Methods of orbit determination.* New York, etc, Wiley (1965) and Malabar (Florida), Krieger (1985)
11	S. Herrick	*Astrodynamics*, Volume 1. London, etc, Van Nostrand Reinhold (1971)
12	R.E. Briggs J.W. Slowey	*An iterative method of orbit determination from three observations of a nearby satellite.* Smith. Inst. Astr. Obs. Special Report 27 (1959)

REFERENCES (concluded)

No.	Author	Title, etc
13	M.T. Lane	A numerical approach to the angles-only initial orbit determination problem. *Adv. Astronaut. Sci.*, **76**, 55-72 (1992)
14	R.H. Gooding	*Universal procedures for conversion of elements to and from position and velocity (unperturbed orbits).* RAE Technical Report 87043 (1987) Also (shortened) in *Celestial Mech.*, **44**, 283-298 (1988)
15	Y. Bard	*Nonlinear parameter estimation.* New York, etc, Academic Press (1974)

UNLIMITED

As far as possible this page should contain only unclassified information. If it is necessary to enter classified information, the box above must be marked to indicate the classification, *eg* Restricted, Confidential or Secret.

1. DRIC Reference (to be added by DRIC)	2. Originator's Reference	3. Agency Reference	4. Report Security Classification/Marking
	DRA TR 93004		UNCLASSIFIED UNLIMITED

5. DRIC Code for Originator	6. Originator (Corporate Author) Name and Location
7673000W	DRA Farnborough, Hampshire, GU14 6TD
5a. Sponsoring Agency's Code	6a. Sponsoring Agency (Contract Authority) Name and Location

7. Title

A New Procedure for Orbit Determination Based on Three Lines of Sight (Angles Only)

7a. (For Translations) Title in Foreign Language

7b. (For Conference Papers) Title, Place and Date of Conference

8. Author 1, Surname, Initials	9a. Author 2	9b. Authors 3,4 ...	10. Date	Pages	Refs
Gooding, R.H.			April 1993		15

11. Contract Number	12. Period	13. Project	14. Other Reference Nos.
			Space 689

15. Distribution statement

(a) Controlled by – Unlimited distribution

(b) Special limitations (if any) –

If it is intended that a copy of this document shall be released overseas refer to DRA Leaflet No. 3 to Supplement 6 of MOD Manual 4

16. Descriptors (Keywords) (Descriptors marked * are selected from TEST)

Orbits*. Celestial mechanics*.

17. Abstract

A new procedure has been developed for the general solution of the minimal angles-only problem in which an orbit is determined from three line-of-sight observations. The basis of the approach is a higher-order Newton correction of the assumed values for two of the unknown ranges, appeal being made to the author's (published) universal solution of Lambert's orbital boundary-value problem.

The new procedure is free of the inherent limitations of the traditional methods of Laplace and Gauss, these methods being outlined in a summary of previous approaches to this classical problem. In particular, the observations are permitted to span several revolutions when the orbit is elliptic; multiple solutions can be obtained; and there is no restriction on the configuration of the three observing sites.

The procedure has been carefully tested, some of the examples being taken from the literature. A number of test problems have been solved that would have failed by existing methods.

F5910/1