A Homotopy Approach to the Solutions of Minimum-Fuel Space-Flight Rendezvous Problems.

by

Gopal Vasudevan

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Aerospace & Ocean Engineering

APPROVED:

Dr.Frederick H. Lutze, Chairman	Dr.Layne T. Watson
Dr.Eugene M. Cliff	Dr.John A. Burns
Dr.Robert W. Walters	Dr.Scott L. Hendricks

February 1989

Blacksburg, Virginia

A Homotopy Approach to the Solutions of Minimum-Fuel Space-Flight Rendezvous Problems.

by

Gopal Vasudevan

Dr.Frederick H. Lutze, Chairman

Aerospace & Ocean Engineering

(ABSTRACT)

A homotopy approach for solving constrained parameter optimization problems is examined. The first order necessary conditions, with the complementarity conditions represented using a technique due to Mangasarian, are solved. The equations are augmented to avoid singularities which occur when the active constraint set changes. The Chow-Yorke algorithm is used to track the homotopy path leading to the solution to the desired problem at the terminal point.

Since the Chow-Yorke algorithm requires a fairly accurate computation of the Jacobian matrix, analytical representation of the system of equations is desired. Consequently, equations obtained using the true anomaly regularization of the governing equations were employed for the above purpose. A homotopy map suited for the space-flight rendezvous problem including a minimum radius constraint is developed, which can naturally deform any initial problem into some other valid desired problem. Several coplanar and non-coplanar solutions for circular and elliptic cases have been presented for the restricted time problem with a minimum radius constraint.

Acknowledgements

The author wishes to express his deep appreciation to his advisor, Dr. Frederick H. Lutze and to Dr. Layne T. Watson for the countless hours of stimulating discussions and for the constant encouragement and assistance provided for the succesful completion of this study.

Thanks are also due to Dr. Eugene M. Cliff and Dr. John A. Burns for their helpful comments and suggestions towards the fulfillment of this endeavor. The author also wishes to thank Dr. Robert W. Walters and Dr. Scott L. Hendricks for being on his committee.

He expresses his appreciation to all his fellow graduate students for their help and assistance.

Table of Contents

List of Illustrations

1.0 Introduction.

Minimum-fuel transfer between orbits in an inverse-square gravitational field is one of the oldest problems in space-flight optimization. One version of this problem assumes a finite number of velocity impulses and coasting arcs to accomplish some space-flight maneuver such as intercept, rendezvous or simply an orbital transfer. Interest in this version stems from the fact that velocity impulses provide a convenient approximation to short thrusting periods. This formulation facilitates modelling the fuel optimization problem for some space-flight maneuvers for solution via a parameter optimization method.

For a given final time, a rendezvous with a non-maneuvering target or an orbital transfer can be accomplished using two impulses. In fact for a circle-to-circle orbit transfer, with a sufficiently long transfer time where the ratio of the final radius to initial radius is less than about 15.6, a particular type of two impulse transfer called a Hohmann transfer is the most fuel efficient.[14, 26, 32, 33] However, when the total time is limited or when the radius ratio becomes greater than 15.6, more impulses may be necessary for optimality. Indeed Hoelker and Silber[21] found bi-elliptic transfers to be fuel-optimal for time open transfers

between orbits with radius ratios in excess of 15.6. A good insight to the general nature of fuel optimal transfers and rendezvous problems can be found in the survey papers by Edelbaum[14] and by Gobetz and Doll.[19]

In applications work, total maneuvering time is often limited and yet fuel efficient transfers must be achieved. It would be of interest to know from the standpoint of parameter optimization, whether or not a fuel-optimal maneuver can be accomplished using a finite number of impulses. It has been shown by Neustadt[36] and by Potter and Stern[41] that for a linearized problem, a minimum-fuel transfer between neighboring orbits can always be done with a number of impulses at most equal to the the number of terminally specified state variables. Thus, if the equations of motion linearized about some mean orbit conditions, such as the Clohessy-Wiltshire[9] equations, are used to describe the problem, then a fuel-optimal rendezvous trajectory can be described with at most four impulses for the coplanar problem, whereas a non-coplanar problem would require not more than six velocity impulses. Unfortunately, not much can be said on the upper limit on the number of impulses for the nonlinear problem, and less if the problem is additionally constrained.

An important version of minimum fuel space-flight rendezvous problems is the Lawden's problem which assumes a variable thrust rocket with constant jet exhaust velocity and unbounded thrust magnitude in a general gravitational field. A comprehensive study of this above problem is due to Lawden[26, 27, 28] who provided a set of necessary conditions for optimality in terms of a *primer vector* . The primer vector is defined as the vector of the costate variables corresponding to the velocity components. For the special case of unbounded thrust, the primer vector theory can be shown to apply to the above mentioned impulsive thrusting model. Lion and Handelsmann[30] later extended the

primer-vector theory to non-optimal trajectories to indicate how time-fixed trajectories could be improved by introducing either additional impulses or by adding initial or final coasts. Using this theory, Prussing[42,43] provided a method of solution for the linearized version of multi-impulse fixed time circle-to-circle rendezvous problem. He linearized the equations of motion about a reference circular orbit, in the vicinity of the terminal orbits, and then solved the boundary-value problem analytically. Glandorf[17] investigated the nonlinear problem for both bounded and impulsive thrusting and included a minimum radius constraint in his analysis. Numerical results for two-impulse fixed-time rendezvous and for one- and two-impulse fixed-time intercept problems were also presented. Chiu[6,7] analyzed the nonlinear time-fixed multi-impulse rendezvous problem. He used the primer vector theory to obtain minimum fuel solutions for circle-to-circle rendezvous problems for both coplanar and a restricted class of non-coplanar cases. He studied the effects of the phase difference and of the ratio of terminal radii on the minimum fuel solutions to the rendezvous problems.

Neustadt[37] showed that any minimum-fuel solution to Lawden's problem can be approximated arbitrarily closely with a finite number of impulses and coasting arcs. An impulse represents a discontinuous change in velocity. Assuming that the trajectory is comprised solely of impulses and coasting arcs, the fuel-optimal rendezvous problem reduces to the selection of an optimal number of impulses in terms of their points of application along the trajectory and their magnitudes and directions. This formulation of the fuel-optimal rendezvous problem makes it particularly suited to solution using parameter optimization techniques.

In order to formulate the problem as a parameter optimization problem or equivalently a non-linear programming problem, one must select a cost function and the parameters

that affect the chosen cost function subject to possible constraints. A formulation of this type is suggested in ref. 31 and is presented here for discussion.

The problem can be characterized by a series of Keplerian coasting arcs separated by velocity impulses. The initial coasting arc must coincide with the current position and orbit of the interceptor, and the final coasting arc must intercept the target. The relative velocity between the target and the interceptor at intercept determine the final velocity impulse required for rendezvous. The sum of the magnitude of this final impulse plus the magnitudes of those which occur between the coasting arcs serve as the cost function to the problem. The parameters which affect this function are those associated with each velocity impulse (e.g. the three components of the impulse), those associated with the time or location of the impulse, (e.g. the change in true anomaly on each Keplerian coasting arc) and the target location (e.g. the change in true anomaly of the target since the start of the maneuver). Constraints are imposed on the problem which enforce the intercept and rendezvous requirement, (i.e., final position and velocity match) and which ensure a feasible solution (e.g. coasting arcs cannot penetrate the Earth or its atmosphere).

The procedure then is to select a specified number of impulses (or coasting arcs) and to solve the related non-linear programming problem using some algorithm.[29] If the algorithm converges to a solution, the solution must be examined to be sure the parameters associated with coasting angles or magnitudes at velocity impulse values are non-zero. If such is not the case, indications are that the number of impulses for an optimal solution is at best one less than the number assumed.

One of the principal problems that was faced in the earlier approach of using a variable-metric algorithm was the need for a good initial guess, in the absence of which

the algorithm either did not converge or converged slowly requiring many iterations. A method which seemed to alleviate the above problem was to use a solution obtained previously corresponding to certain set of parameters as an initial guess for the solution corresponding to adjacent values of the parameters. Often even this procedure also failed to provide the required convergence.

Thus alternate ways of seeking solutions, based on the above idea of continuation, to the orbital rendezvous problems are investigated. To facilitate such a study, the orbital rendezvous problem as described above is completely reformulated in terms of regularized variables[5,48] with appropriate variables describing an impulse. The above formulation provides a completely analytic description for all quantities such as the characteristic velocity, time to rendezvous, the minimum radius encountered on each subarc of the rendezvous trajectory etc. and their derivatives.

Continuation methods or more general homotopy methods have been used in the past for a variety of problems, and were originally used in an effort to globalize Newton's method.[15] More recently these methods have found a myriad of applications such as in the solution of two-point boundary value problems,[45] in step size control,[11] in the solution to fixed point problems and nonlinear equations.[8,12,22,47,52] Several engineering applications have been detailed in Ref 51. We now seek to develop a constrained parameter optimization algorithm based on homotopy methods, specifically the algorithm due to Chow, Mallet-Paret, and Yorke,[8] or at least to provide a more systematic method to obtain solutions corresponding to set of system constants. We have applied the above algorithm to one simple example to illustrate the method and also used it to obtain solutions to the fuel-optimal orbital rendezvous problems for both

coplanar and non-coplanar cases. We shall discuss continuation and homotopy methods in greater detail in the next two chapters.

2.0 Continuation Methods

2.1 Introduction.

Iterative techniques such as Newton's methods and its variants form an important class of methods to solve non-linear algebraic equations. These methods are locally convergent and are based on the contraction principle. In other words, for such algorithms to yield a solution, a starting point 'sufficiently close' to the actual solution must be supplied, where the term 'sufficiently close' depends on the problem solved and is not known *a priori.* For low order problems, or for problems which are well understood, providing an initial guess is not difficult or computationally expensive. But with more ambitious problems, even supplying an initial value for the algorithm to start is a formidable task, and often some other more robust methods have to be employed to obtain the starting value. Clearly, this need for a good starting value imposes a serious restriction for a Newton-type iterative method.

It is for these reasons (amongst others) that homotopy methods, continuation methods, or imbedding methods are employed in the hope of dispensing with the requirement of a good starting guess. Numerical continuation methods have been widely used to obtain families of solutions with respect to some parameter (intrinsic or otherwise). The principal idea of continuation is to use the solution of a problem corresponding to a particular value of a parameter ε as a starting point to obtain the solution corresponding to nearby value of the parameter $\varepsilon + \Delta\varepsilon$. The assumption behind this is that if there are small variations in the problem constants, the problem characteristics are essentially unchanged, and the solution to the new problem should be in the neighborhood of the solution to the initial problem.

A principal development in continuation methods, due to Davidenko,[10] was the proposal to differentiate the homotopy equation with respect to the parameter. Together with the known solution for some value of the parameter, it gives an initial value problem. Yakovlev[59] examined this method from a theoretical standpoint to determine when the resulting Cauchy problem has a solution. He also looked at some ways of introducing the continuation parameter so that if certain conditions are met the resulting continuation equation would have unique solutions for all values of the parameter in the range [0, 1]. In the usual applications of continuation methods, the principal assumption is that the solution to the continuation equation is a continuous single-valued function of the imbedding parameter.

A related idea in the specialized context of boundary value problems is the method of *invariant imbedding*. This technique involves using some parameter to generate a family of problems, encompassing the desired problem, so that it retains certain intrinsic properties of the original problem. The idea is to solve a stable initial value problem

starting from the solution to a trivial, maybe degenerate, member of the family of problems to obtain the solution to the desired, possibly unstable, boundary value problem. In its original applications, physical properties such as the size of an object, length of a beam or an interval provided the imbedding parameter. A comprehensive introduction to this area may be found in Scott,[49] or Bellman and Wing.[3]

The basic idea behind the continuation methods or more general homotopy methods is to construct a path joining the solutions to two problems, one of which may be a trivial problem which resembles in some sense the problem to be solved, and the solution to which is known or at least is easily obtainable. The second problem is the problem of interest. For example, if $f_0(x) = 0$ is the trivial problem and $F(x) = 0$ is the problem of interest, then one can construct the homotopy path as the solutions of:

$$\sigma F(x) + (1 - \sigma) f_0(x) = 0, \qquad [2.1.1]$$

for $0 \leq \sigma \leq 1$. Then using some curve following technique, one can obtain the solution to the desired problem (at $\sigma = 1$) starting from the solution to the trivial problem (at $\sigma = 0$).

Philosophically, the homotopy methods differ from classical continuation methods in that the constructed homotopy map is viewed as a systematic deformation of the trivial problem to obtain the desired problem. Thus, the parameter σ is viewed merely as some variable aiding the deformation process. Thus in following the homotopy path, the solutions to the homotopy equation may not always vary monotonically with respect to the homotopy parameter σ, in contrast to the basic assumption of continuation methods. The basic theory behind homotopy methods come from algebraic and/or differential topology.

The first computational algorithm based on the theory of simplicial approximations for continuous mappings (with roots in algebraic topology) was proposed by Scarf,[47] and has since been improved by Eaves,[12,13] Saigal[46] and many others. An excellent historical survey of the development of the simplicial continuation algorithm is provided by Allgower and Georg[1]. They have also included an elementary simplicial continuation algorithm in their paper.

On a different note, based on the differentiability assumption, Kellogg, Li, and Yorke[23] provided an algorithm for the fixed point problem. Similar algorithms were proposed by Chow, Mallet-Paret and Yorke,[8] Keller,[22]Klopfenstein,[24]Kubicek,[25] Watson,[52] and others. An interesting historical perspective on the development of the continuation methods may be obtained from Wacker[50] and on homotopy methods from Watson.[54]

2.2 *Example: Solution to Kepler's Equation*

Let us consider now an example to demonstrate a simple continuation method. Consider the solution to Kepler's equation, i.e., given the mean anomaly M, and the eccentricity of the orbit e, solve for the eccentric anomaly E.

$$E - e \sin(E) - M = 0, \quad \text{for } e < 1. \qquad [2.2.1]$$

One can set up a continuation algorithm using either M or e as the continuation parameter.

Let e be the continuation parameter. We are interested in finding the solution to the problem for a given value of M, and for some value of $e = e_0 < 1$. To build a continuation map, we note that when $e = 0$ the solution to our problem is $E = M$. Using this as our trivial problem, we can construct a continuation map as

$$E - \sigma e_0 \sin(E) - M, \qquad [2.2.2]$$

so that when $\sigma = 0$, we have the problem $E - M = 0$, and at $\sigma = 1$, we have the original problem. Note also that as the continuation parameter σ varies from 0 to 1, we obtain a family of solutions corresponding to e varying from 0 to e_0. The solutions to the above equation for any value of σ can be obtained using any curve tracking procedure, typically a procedure based on the solution to an implicit ordinary differential equation.

A Simple Continuation Method

We shall describe now a simple procedure to track the zero curve of the continuation map. Differentiating equation [2.2.2] with respect to the continuation parameter σ, we obtain:

$$(1 - \sigma e_0 \cos(E)) \frac{dE}{d\sigma} - e_0 \sin(E) = 0. \qquad [2.2.3]$$

The essential idea is to solve the initial value problem obtained above. One may use a Runge-Kutta type method, or an Adams-Bashforth type predictor-corrector method, or any other favorite algorithm. We shall describe a simple predictor-corrector method to illustrate the procedure. Choose a small step $\Delta\sigma$ for σ and from equation [2.2.3] compute a change in E. The predicted value of $E = E^{\text{pred.}}$ is then used as an initial guess in a Newton iteration, i.e.,

$$E^{k+1} = E^k - \frac{E^k - \sigma^{new} e_0 \sin(E^k) - M}{1 - \sigma^{new} e_0 \cos(E^k)}, \qquad [2.2.4]$$

for $k = 1, \ldots$. to obtain the correct solution to equation [2.2.2] at $\sigma^{new} = \sigma^{old} + \Delta\sigma$, where $E^1 = E^{pred.}$. This sequence of predictor and corrector steps, is repeated starting from $E = M$ at $\sigma = 0$ until $\sigma = 1$. The solution thus obtained gives the value of $E = E^*$ such that $E^* - e_0 \sin(E^*) = M$. Figure 1 gives a plot of the eccentric anomaly versus the eccentricity.

Similarly we could have chosen the mean anomaly itself as the continuation parameter with the continuation map as:

$$E - e_0 \sin(E) - \sigma M. \qquad [2.2.5]$$

Figure 2 shows the variation of the eccentric anomaly with respect to the mean anomaly for several values of the eccentricity. For both examples, the step $\Delta\sigma$ was arbitrarily picked to be 10^{-2}, and the correction to the solution curve was done with an absolute tolerance $|E^{k+1} - E^k| \leq 10^{-8}$. The number of correction steps to attain this accuracy was at most 5 for each step.

2.3 *A Continuation Approach to Constrained Parameter Optimization*

The above procedure illustrates the basic idea behind continuation methods. This continuation idea can now be applied to solve a parameter optimization problem by finding the roots of the Kuhn-Tucker equations.

Consider a constrained optimization problem **p**. Find $\bar{x}$ such that

$$C(\bar{x}) = \min_{x \in X} C(x); \quad X = \{ x \mid x \in \mathcal{R}^n, g(x) \geq 0, \quad h(x) = 0 \}, \qquad [2.3.1]$$

where $C: \mathcal{R}^n \to \mathcal{R}$, $g: \mathcal{R}^n \to \mathcal{R}^{ni}$ and $h: \mathcal{R}^n \to \mathcal{R}^{ne}$ are twice continuously differentiable functions.

An appropriate continuation method which applies to this problem is attributed to Gfrerer, Guddat and Wacker.[16] Required, however, are the following two assumptions:

A1 : The above problem can be imbedded in a class of problems $\mathbf{P}(\varepsilon)$ such that for some value ε_1, $\mathbf{P}(\varepsilon_1) = \mathbf{p}$.

$$\begin{aligned} &C(\bar{x}, \varepsilon) = \min_{x \in X} C(x, \varepsilon); \\ &\bar{x} \in X = \{ x \mid x \in \mathcal{R}^n, g(x, \varepsilon) \geq 0, \quad h(x, \varepsilon) = 0 \}. \end{aligned} \qquad [2.3.2]$$

A2 : For some value of the parameter $\varepsilon = \varepsilon_0$, let $\mathbf{P}(\varepsilon_0)$ have a solution $\bar{x}$ so that $\bar{x}$ satisfies the first order necessary conditions, i.e., the Kuhn-Tucker conditions.

The generalized Kuhn-Tucker conditions or equivalently the Fritz-John conditions[32] are:

If $\bar{x}$ is a solution to [2.3.2] for a certain value of the parameter ε then there exists a scalar $\bar{\mu}_0$ and vectors $\bar{\lambda} \in \mathcal{R}^{ne}$ and $\bar{\mu} \in \mathcal{R}^{ni}$ such that:

$$\bar{\mu}_0 \nabla C(\bar{x}, \varepsilon) - \bar{\lambda}^T \nabla h(\bar{x}, \varepsilon) - \bar{\mu}^T \nabla g(\bar{x}, \varepsilon) = 0, \qquad [2.3.3]$$

$$h(\bar{x}, \varepsilon) = 0, \qquad [2.3.4]$$

$$\bar{\mu}^T g(\bar{x}, \varepsilon) = 0, \qquad [2.3.5]$$

$$g(\bar{x}, \varepsilon) \geq 0, \quad \bar{\mu} \geq 0, \qquad [2.3.6]$$

$$\bar{\mu}_0 \geq 0, \quad [\bar{\lambda}, \bar{\mu}, \bar{\mu}_0] \neq 0. \qquad [2.3.7]$$

where $\neq 0$ means that some but not all of the terms may be zero. We normally assume $\bar{\mu}_0 = 1$ (the Kuhn-Tucker conditions), hoping that the Kuhn-Tucker constraint qualification[32] will be satisfied at $\bar{x}$.

Usually an appropriate linear scaling of the parameter ε can be done such that:

$$\varepsilon(\sigma) = \sigma \varepsilon_1 + (1 - \sigma) \varepsilon_0.$$

Solution Procedure

A continuation method can now be applied to extend the solution of the problem $\mathbf{P}(\sigma = 0)$ to the solution of the problem $\mathbf{P}(\sigma = 1)$, by continuously finding the roots of the equality expressions. In other words, at any given step the inactive inequality

constraints are not considered. This involves keeping track of the active constraints at every step.

Let **I** represent the index set of active constraints, i.e.,

$$\mathbf{I} = \{i \mid g_i(\bar{x}, \sigma) = 0\}, \qquad i = 1, \ldots, \text{ni}$$

where g is the inequality constraint. Thus the Kuhn-Tucker equations solved by the continuation method can be written as follows:

$$\nabla C(\bar{x}, \sigma) - \bar{\lambda}^T \nabla h(\bar{x}, \sigma) - \bar{\mu}^T \nabla g(\bar{x}, \sigma) = 0, \qquad [2.3.8]$$

$$h(\bar{x}, \sigma) = 0, \qquad [2.3.9]$$

$$g_i(\bar{x}, \sigma) = 0, \qquad \text{for } i \in \mathbf{I} \qquad [2.3.10]$$

$$\bar{\mu}_i = 0, \qquad \text{for } i \notin \mathbf{I} \qquad [2.3.11]$$

One can now write a predictor scheme as follows: Let

$$\mathscr{L}(x, \lambda, \mu, \sigma) = C(x, \sigma) - \lambda^T h(x, \sigma) - \mu^T g(x, \sigma),$$

then the predictor equation is given by:

$$\begin{bmatrix} \nabla^2 \mathscr{L}(\bar{x}^k, \bar{\lambda}^k, \bar{\mu}^k, \sigma^k) & -\nabla h(\bar{x}^k, \sigma^k) & -\nabla g_I(\bar{x}^k, \sigma^k) \\ \nabla h^T(\bar{x}^k, \sigma^k) & 0 & 0 \\ \nabla g_I^T(\bar{x}^k, \sigma^k) & 0 & 0 \end{bmatrix} \begin{Bmatrix} x' \\ \lambda' \\ \mu'_i \end{Bmatrix} = - \begin{Bmatrix} [\nabla \mathscr{L}(\bar{x}^k, \bar{\lambda}^k, \bar{\mu}^k, \sigma^k)]' \\ h'(\bar{x}^k, \sigma^k) \\ g'_I(\bar{x}^k, \sigma^k) \end{Bmatrix}, \qquad [2.3.12]$$

where ()′ denotes the derivative with respect to the parameter and $\{\bar{x}^k, \bar{\lambda}^k, \bar{\mu}^k\}$ is the solution to the system of equations [2.3.8]-[2.3.11] corresponding to $\sigma = \sigma^k$.

The predicted value can be computed by taking a simple Euler step. Since the predicted value is the solution to a linearized problem, the prediction may not lie on the curve. Consequently, a correction step must be made. The correction step or restoration step must be made in such a way so as not to nullify the prediction The simplest way to achieve this is to freeze the value of the continuation parameter $\sigma = \sigma^{k+1}$ and correct only for the Kuhn-Tucker triple $\{x, \lambda, \mu\}$. A Newton type method is used to perform the correction:

$$\begin{bmatrix} \nabla^2 \mathscr{L}(x, \lambda, \mu, \sigma^{k+1}) & -\nabla h(x, \sigma^{k+1}) & -\nabla g_I(x, \sigma^{k+1}) \\ \nabla h^T(x, \sigma^{k+1}) & 0 & 0 \\ \nabla g_I^T(x, \sigma^{k+1}) & 0 & 0 \end{bmatrix} \begin{Bmatrix} \Delta x \\ \Delta \lambda \\ \Delta \mu_I \end{Bmatrix} = - \begin{Bmatrix} [\nabla \mathscr{L}(x, \lambda, \mu, \sigma^{k+1})] \\ h(x, \sigma^{k+1}) \\ g_I(x, \sigma^{k+1}) \end{Bmatrix}. \quad [2.3.13]$$

At the end of a predictor-corrector cycle, one must check the constraint set to see if:

i. $g_i(\bar{x}^{k+1}, \sigma^{k+1}) \leq 0$, for $i \notin \mathbf{I}$,

ii. $\mu_i \leq 0$, for $i \in \mathbf{I}$,

where $\bar{x}^{k+1}$ is the solution obtained from [2.3.13]. If either of the above conditions hold true then the constraint set must be modified. To do this, the earliest location $\tilde{\sigma}$ at which

i. $g_i(x, \tilde{\sigma}) = 0$, for some $i \notin \mathbf{I}$ or

ii. $\mu_i = 0$, for $i \in \mathbf{I}$,

holds true must be determined accurately.

Suppose the inequality constraint $g_i(x, \sigma)$, $i \notin \mathbf{I}$, becomes active, i.e., $g_i(\bar{x}^{k+1}, \sigma^{k+1}) \leq 0$. The location, $\tilde{\sigma}$, may be computed using a chord method, since it is bracketed by σ^k and σ^{k+1}. Thus, on the curve of solutions

$$g_i(\bar{x}^k, \sigma^k) > 0, \quad \text{and } g_i(\bar{x}^{k+1}, \sigma^{k+1}) < 0,$$

where $i \notin \mathbf{I}$. The following simple chord method was used to obtain the precise location of active constraint set changes for the simple parameter optimization example given below.

A predictor set of values are computed at:

$$\tilde{\sigma} = \sigma^k - \frac{g_i(\bar{x}^k, \sigma^k)\,(\sigma^{k+1} - \sigma^k)}{g_i(\bar{x}^{k+1}, \sigma^{k+1}) - g_i(\bar{x}^k, \sigma^k)}. \qquad [2.3.14]$$

Since the Jacobian matrix of the system of equations is already available at σ^{k+1}, it is a simple matter to obtain a predicted set of values for the Kuhn-Tucker triple at $\tilde{\sigma}$. These predicted values of $\{x, \lambda, \mu\}$ are then used in the corrector cycle to obtain the solution at $\tilde{\sigma}$. Replace σ^k or σ^{k+1} with $\tilde{\sigma}$, depending on the value of $g_i(x, \tilde{\sigma})$. Compute a new $\tilde{\sigma}$, and the Jacobian matrix at $\tilde{\sigma}$, and repeat the above process of sequential predictor and corrector cycles until $|g_i(x, \tilde{\sigma})| \leq$ tol. For each of the corrector cycles we may use the same Jacobian matrix used in the predictor computation.

The above procedure may be modified appropriately, if μ_i, $i \in \mathbf{I}$, changes sign. Of course, this is only a prototype procedure and is quiet inefficient. More sophisticated algorithms may be employed for this purpose[55].

The constraint set is updated and the classical continuation procedure is resumed.

2.4 *Example*

Let us consider a simple example to examine this method.

$$P(\bar{x}, \sigma) = \min_{x \in X} \; x_1^2 + x_1 x_2 + (\alpha - 1) x_2^2$$
$$\bar{x} \in X = \{ x \mid x \in \mathcal{R}^2, \; g(x) \geq 0 \},$$

where the inequality constraints are defined as

$$g_1(x) \equiv 2x_2 - x_1 \geq 0,$$

$$g_2(x) \equiv 2x_1 - x_2 \geq 0,$$

$$g_3(x) \equiv x_1^2 + x_2^2 - 1 \geq 0,$$

$$g_4(x) \equiv 2 - x_1^2 - x_2^2 \geq 0,$$

and $\alpha = \sigma \alpha_1 + (1 - \sigma) \alpha_0$, $\alpha_0 = 0$, $\alpha_1 = 3$.

For this problem, the Lagrangian is given by

$$\mathcal{L} = x_1^2 + x_1 x_2 + (\alpha - 1) x_2^2 - \mu_1 (2x_2 - x_1) - \mu_2 (2x_1 - x_2) - \mu_3 (x_1^2 + x_2^2 - 1) - \mu_4 (2 - x_1^2 - x_2^2).$$

Let $\alpha_0 = 0$ be the starting problem, for which we know the solution to be:

$$x_1 = 0.6324555, \quad x_2 = 1.2649111, \; \mu_1 = 0.0, \; \mu_2 = 1.3914022, \; \mu_3 = 0.0, \; \mu_4 = 0.2 .$$

The feasible region for this problem is as shown in figure 3. This is a simple example for which the solution is known analytically for all values of α. It can be easily shown that depending on the value of α each of the corners is a locally optimal solution. For this problem for $\alpha \leq 0.25$ the top left corner is the optimal solution. At $\alpha = 0.25$ all feasible points on the constraint g_1 are solutions. For $0.25 \leq \alpha \leq 2.75$ the bottom left corner provides a locally optimal solution, and for $1.25 \leq \alpha$ the bottom right corner provides a locally optimal solution. Note that there are multiple solutions for the values of α for $1.25 \leq \alpha \leq 2.75$.

When $\alpha = 0$, the constraints g_1 and g_3 are inactive. Using the above method we note that when $\alpha = 0.25$, g_4 constraint becomes inactive, i.e., $\mu_4 = 0$. On restarting with the solution at $\alpha = 0.25$ it is seen that the Jacobian matrix thus obtained is singular and consequently the method fails to proceed. Physically what has happened is that at this value of $\alpha = 0.25$, any point on the constraint g_2 furnishes a minimum, i.e., we have an improper minimum.

The above simple example reveals a drawback of the classical continuation methods. The main assumption of these algorithms is the requirement of a non-singular Jacobian matrix over the entire continuation path.[59] In other words, classical continuation methods require the solutions to be monotonic with respect to the parameter.

Following an idea due to Chow, Mallet-Paret, and Yorke,[8] Keller,[22] and Watson,[52] we choose the arc-length s as the continuation parameter instead of using the parameter α or equivalently σ, and treat σ as simply another variable. We then find a unit vector in the null space of the Jacobian matrix with respect to $\{ x, \lambda, \mu, \sigma \}$, and then determine the predicted value of these variables based on the previous step and direction of motion. In other words:

i. Find a non-zero vector $\{z\}$ in the null space of the Jacobian matrix. This is nothing more than finding a tangent vector to the curve in $n+1$ dimensional space, where n is the dimension of the Kuhn-Tucker triple $\{x, \lambda, \mu\}$.

$$\begin{bmatrix} \nabla^2 \mathcal{L}(\bar{x}^k, \bar{\lambda}^k, \bar{\mu}^k, \sigma^k) & -\nabla h(\bar{x}^k, \sigma^k) & -\nabla g_I(\bar{x}^k, \sigma^k) & [\nabla \mathcal{L}(\bar{x}^k, \bar{\lambda}^k, \bar{\mu}^k, \sigma^k)]_\sigma \\ \nabla h^T(\bar{x}^k, \sigma^k) & 0 & 0 & h_\sigma(\bar{x}^k, \sigma^k) \\ \nabla g_I^T(\bar{x}^k, \sigma^k) & 0 & 0 & g_{I\sigma}(\bar{x}^k, \sigma^k) \end{bmatrix} z = 0, \quad [2.3.14]$$

where $z = \{x', \lambda', \mu'_I, \sigma'\}$ and $()'$ denotes the derivative with respect to the arc-length and $()_\sigma$ denotes the derivative with respect to σ.

ii. Normalize this vector.

$$\begin{Bmatrix} x' \\ \lambda' \\ \mu'_I \\ \sigma' \end{Bmatrix} = \pm \frac{z}{\| z \|_2}. \quad [2.3.15]$$

There is a sign ambiguity that results in this normalization, which relates to determining the direction of this unit vector. The correct sign is fixed by choosing it to be such that the unit tangent vector makes an acute angle with the tangent vector at the previous step. At the start of the algorithm, the sign is chosen to provide a positive change for σ. Allgower and Georg[1] propose a different method to choose the sign of the normalization, with the same effect.

With this modification, if we proceed with our example, then as shown in Table 1, at $\alpha = 0.25$ after the constraint set update, the method proceeds with no change in α until $x_1 = 0.4472$ and $x_2 = 0.8944$. At this point the constraint g_3 becomes active. After re-updating the constraint set, the method proceeds until $\alpha = 2.75$ at which point constraint g_2 becomes inactive. The next step shows a negative change in the value of α. Unlike the classical continuation methods, here we now continue on in the hope that

the curve would turn again. Indeed, at $\alpha = 1.25$, the constraint g_1 becomes active, and the curve turns again. From here on the α increases monotonically.

Table 1. Simple continuation path including singular points for the above example

α	x_1 , x_2		μ_1 , μ_2, μ_3, μ_4				Index set	Comments
0.0 0.0415 0.1658 0.2124 0.2357	0.6325	1.2649	0.0	1.3914 1.370 1.308 1.284 1.272	0.0	0.2000 0.1668 0.0674 0.0301 0.0114	2, 4	
0.2500	0.6325	1.2649	0.0	1.2649	0.0	0.0	2	critical point
0.25	0.5851 0.5218 0.4899	1.1704 1.0437 0.9799	0.0	1.1704 1.0437 0.9799	0.0	0.0	2	
0.25	0.4472	0.8944	0.0	0.8944	0.0	0.0	2	critical point
0.5161 0.6376 1.002 2.042 2.562 2.692 2.741	0.4472	0.8944	0.0	0.7992 0.7557 0.6253 0.2533 0.0673 0.0208 0.0034	0.2129 0.3101 0.6017 1.434 1.849 1.953 1.992	0.0	2, 3	
2.750	0.4472	0.8944	0.0	0.0000	2.000	0.0	3	turning point
2.730 2.687 1.941 1.512 1.350	0.4529 0.4658 0.7275 0.8481 0.8790	0.8916 0.8849 0.6861 0.5298 0.4769	0.0	0.0	1.984 1.950 1.472 1.312 1.271	0.0	3	
1.250	0.8944	0.4472	0.0	0.0	1.250	0.0	3	turning point
1.578 1.731 2.189 3.0	0.8944	0.4472	0.1175 0.1721 0.3358 0.6261	0.0	1.316 1.346 1.438 1.600	0.0	1, 3	

2.5 *Poore and Al-Hassan's Algorithm*

A related continuation based algorithm is due to Poore and Al-Hassan.[40] This method differs from the previous continuation algorithm in that the solution is not obtained for a range of values of the parameter of interest. The technique is essentially a refinement of the classical SUMT (sequential unconstrained minimization technique). The procedure is to successively minimize an auxiliary function P , which contains the constraints in so-called penalty barrier functions collectively weighted by some constant r, i.e.,

$$\begin{aligned} &\underset{x \in X}{\text{Min}}\ P = C(x) + \frac{1}{2r} h^T(x)\, h(x) - r \sum_{i=0}^{ni} \ln(g_i(x)) \\ &X = \{ x \mid g(x) > 0 \}, \end{aligned} \qquad [2.5.1]$$

where r is the penalty constant.

Ideally, as the penalty constant tends to zero the unconstrained minimum of P approaches the solution to [2.3.1]. However, the sequential unconstrained minimization techniques are wrought with convergence problems owing to the eventual ill-conditioning of the Hessian of P as the parameter r becomes small.

To remove such a singularity Poore and Al-Hassan propose an algorithm based on expanding the gradient of P:

$$\nabla P(x, r) = \nabla C(x) + \frac{h^T(x)}{r} \nabla h(x) - \sum_{i=1}^{ni} \frac{r}{g_i(x)} \nabla g_i(x) = 0. \qquad [2.5.2]$$

Thus $(\tilde{x}, \tilde{r})$, $\tilde{x} \in X$, $\tilde{r} > 0$, is a solution to the above equation if and only if it is a solution to

$$\nabla C(x) - \lambda^T \nabla h(x) - \mu^T \nabla g(x) = 0, \qquad [2.5.3]$$

where λ and μ are defined by:

$$h(x) + r\lambda = 0, \qquad [2.5.4]$$

$$\mu_i g_i(x) - r = 0, \quad i = 1, \ldots, \text{ni}. \qquad [2.5.5]$$

In their algorithm, sequential minimization is stopped at some small value of the penalty constant $r = \tilde{r}$. The solution $(\tilde{x}, \tilde{r})$, $\tilde{x} \in X$, $\tilde{r} > 0$, obtained must satisfy [2.5.2]. A simple homotopy map is then constructed to join the solution of the auxiliary problem at this value of the penalty constant $\tilde{r}$ to the solution to the actual problem with the penalty constant r serving as the continuation parameter. A simple linear scaling of r is used as follows

$$r = (1 - \sigma)\tilde{r}.$$

Poore and Al-Hassan used an additional equation to normalize the multipliers. Their homotopy equations are given by:

$$\mu_0 \nabla C(x) - \lambda^T \nabla h(x) - \mu^T \nabla g(x) - \mu_0 (1 - \sigma) \nabla P(\tilde{x}, \tilde{r}) = 0, \qquad [2.5.6]$$

$$\mu_0 h + \tilde{r}(1 - \sigma)\lambda = 0, \qquad [2.5.7]$$

$$\mu g - \tilde{r}(1 - \sigma) = 0. \qquad [2.5.8]$$

$$\mu_0^2 + \|\lambda\|_2^2 + \|\mu\|_2^2 - \beta_0^2 = 0, \qquad [2.5.9]$$

where $\beta_0 = \sqrt{1 + \|\lambda\|_2^2 + \|\mu\|_2^2}$, is held constant at its initial value. A curve tracking procedure is then used to obtain the solution to Problem P [2.3.1] starting from $\sigma = 0$, with $x = \tilde{x}$, $\lambda = \frac{-h(\tilde{x})}{\tilde{r}}$, $\mu = \frac{\tilde{r}}{g(\tilde{x})}$, and $\mu_0 = 1$.

The unconstrained minimization must be an interior method, i.e., the inequality constraints should never be violated. This involves modifying the convergence or stopping criterion so that the search is always confined to the feasible region. For this purpose one may use a modified Armijo criterion as follows: for a fixed step $s > 0$, $\gamma \in (0, \frac{1}{2})$, and $\beta \in (0, 1)$ find an integer $m = 0, 1, \ldots$ such that

$$P(x_k) - P(x_k + \beta^m s\, d_k) \geq -\gamma\, \beta^m s\, \nabla P^T(x_k)\, d_k,$$

where $\alpha_k = \beta^{m_k} s$ should be such that

$$g_i(x_k + \alpha_k d_k) \geq 0, \qquad i = 1, \ldots, \text{ni}.$$

d_k is the search direction, and β is chosen initially depending on the confidence placed on the stepsize s.

As with any algorithm, this one is also beset with certain problems. The main drawback stems from the assumption made that the unconstrained minimum to the auxiliary problem is in the neighborhood of the solution to [2.3.1] in the sense of the implicit function theorem. For the problems with which we tested this algorithm, there were instances, when the curve simply stopped, thus indicating that there is no curve joining the solutions to the two problems. In other words, the solution to the auxiliary problem corresponding to the penalty constant $r = \tilde{r}$ is not close to the actual solution. One may have to further decrease the value of r. So the question arises of how small the penalty parameter must be made before the continuation procedure can be successfully applied.

Of course, one must also bear in mind the inherent problems with the penalty barrier function method when the values of the penalty parameter are small.

2.6 Remarks Concerning Continuation Methods.

The parameter optimization example given above is a very simple illustration of what can go wrong with a simple continuation method. Also there exist problems where the curve bifurcates. Under such circumstances, one must locate the bifurcation point and then determine which of the branches to follow, generally by using the second order necessary conditions.

There exists another possibility, i.e., the curve simply comes to a stop even before the desired value of the parameter is obtained. This indicates that there is no curve connecting the solutions of the two problems. These possible difficulties in obtaining a solution indicate that a more sophisticated method is needed to achieve the algorithm we are hoping for: one which starts with the solution to some trivial problem and continuously deforms it to obtain the solution to the related desired problem.

3.0 Homotopy Methods

3.1 Introduction

Continuation methods as described earlier have an inherent elegance in their approach to obtain families of solutions. Depending on some parameter of interest one follows a path of solutions to some meaningful problems. However, as explained previously, classical continuation methods are based solely on the premise that the solutions are single valued with respect to the parameter of interest. As shown earlier, this may not always be the case. Even so it may still be possible to follow a curve of solutions by using the arc-length as the continuation parameter. In the event the curve turns, the curve tracking procedure can be continued with the hope that eventually it will reach its destination. Often times step sizes have to be regulated so that the curve tracking procedure does not loop or jump from the desired path to a neighboring undesired path. In spite of the use of arc-length as the independent variable, it is still not guaranteed that a solution exists for all values of the parameter of interest, or the curve may bifurcate. Thus, a means is required which would enable one to ensure at least the existence of a

path. In the latter event where the curve bifurcates, one is forced to look into alternate ways of finding the bifurcating branches, choose the appropriate branch and follow it until the end.

Homotopy methods, based on the parameterized version of Sard's theorem,[8, 20, 52] are designed to remedy the above mentioned shortcomings of continuation.

3.2 *Chow-Yorke Algorithm*

In order to fully appreciate the significance of the Chow-Yorke algorithm, it is first necessary to introduce a few definitions and theorems. First we provide a few definitions taken from algebraic topology and differential geometry:

Definition Let $U,\ Y \subset \mathcal{R}^n$ and $g_a : U \to Y$ and $f : U \to Y$. $g_a(z)$ and $f(z)$ are *homotopic* if there exists a continuous map $\rho : [0, 1] \times U \to Y$ such that $\rho(0, z) = g_a(z)$ and $\rho(1, z) = f(z)$, where $z \in U$. ρ is called a *homotopy* between $g_a(z)$ and $f(z)$.

Definition Let $U,\ V \subset \mathcal{R}^n$ be open sets and $\rho : U \times [0, 1) \times V \to \mathcal{R}^n$ be a C^2 map. ρ is said to be transversal to zero if the Jacobian matrix $D\rho$ has full rank on the set $\rho^{-1}(0) = \{(a,\ \sigma,\ z) |\ \rho(a, \sigma, z) = 0,\ a \in U,\ \sigma \in [0, 1),\ z \in V\}$.

Although the definition of transversality stated here is merely a consequence of the definition stated in the works of Guillemin and Pollack[20] or Bruce and Giblin,[4] it suffices for our purpose to indicate the non-singularity of the Jacobian matrix $D\rho$.

Parameterized Sard's Theorem[8, 52]

If $\rho(a, \sigma, z)$ *is transversal to zero, then for almost all* $a \in U$, *the map*

$$\rho_a(\sigma, z) = \rho(a, \sigma, z)$$

is also transversal to zero, i.e., with probability one the Jacobian matrix $D\rho_a(\sigma, z)$ *has full rank on* $\rho_a^{-1}(0)$.

Thus if $\rho(a, \sigma, z)$ is a homotopy between a simple map $g_a(z) = \rho(a, 0, z)$ and the map of interest $f(z) = \rho(a, 1, z)$, and if ρ is transversal to zero, then $\rho_a(\sigma, z)$ (with the vector a fixed) is also transversal to zero for almost all choices of a.

Implicit Function Theorem[38]

Let $\rho_a : D \subset \mathcal{R}^{n+1} \to \mathcal{R}^n$ *be a* C^2 *map defined on a neighborhood of* $(y_{1_0}, \dots, y_{k_0}, \dots, y_{n+1_0}) \in D$ *with* $\rho_a(y_{1_0}, \dots, y_{k_0}, \dots, y_{n+1_0}) = 0$. *If the* $n \times n$ *Jacobian matrix* $D\rho_a(y_{1_0}, \dots, y_{k-1_0}, y_{k+1_0}, \dots, y_{n+1_0})$ *is non-singular at* $(y_{1_0}, \dots, y_{k-1_0}, y_{k_0}, y_{k+1_0}, \dots, y_{n+1_0})$, *then there exist open neighborhoods, A of* y_{k_0} *and B of* $\{y_{1_0}, \dots, y_{k-1_0}, y_{k+1_0}, \dots, y_{n+1_0}\}$ *such that for all* $y_k \in A$, *there is a unique point* $\varphi_i(y_k)$, $i = 1, \dots, n$, *in B, with*

$$\rho_a(\varphi_1(y_k), \dots, \varphi_{k-1}(y_k), y_k, \varphi_k(y_k), \dots, \varphi_n(y_k)) = 0.$$

Furthermore, the map $\varphi : A \to \mathcal{R}^n$ *is* C^2.

For our problem, let $y = \{\sigma, z\}$. Thus, $y_i, i = 1, \dots, n+1;\ i \neq k$, can be uniquely parameterized in terms of y_k in an open neighborhood of $(y_{1_0}, \dots, y_{n+1_0})$. Note that no

distinction has been placed on either σ or z_i, $i = 1, \ldots, n$. The import of this is that there is a unique smooth, non-bifurcating, non-stopping curve $y_i = \varphi(y_k)$, $i \neq k$, passing through $(y_{1_0}, \ldots, y_{n+1_0})$ within some open neighborhood of this point.

The Chow-Yorke[8] algorithm essentially uses the implicit function theorem in conjunction with the parameterized version of Sard's theorem. Since Sard's theorem guarantees a non-singular Jacobian matrix for the homotopy $\rho_a(\sigma, z) : [0, 1) \times V \to \mathcal{R}^n$, for almost all choices of the fixed parameter vector a, the implicit function theorem consequently provides a smooth, non-bifurcating, and non-stopping curve. Thus, any such curve lying in the region between the hyperplanes $\sigma = 0$ and $\sigma = 1$ would have to do one of the following:

i. start at $\sigma = 0$ and cross $\sigma = 1$ hyperplane,
ii. start at $\sigma = 0$, wander and return to $\sigma = 0$ hyperplane,
iii. start at $\sigma = 1$, wander and return to $\sigma = 1$ hyperplane,
iv. form a closed curve wholly within the region flanked by $\sigma = 0$ and $\sigma = 1$ hyperplanes,
v. start at $\sigma = 0$ and go off to infinity within the region bounded by $\sigma = 0$ and $\sigma = 1$ hyperplanes,
vi. start at $\sigma = 1$ and go off to infinity within the region bounded by $\sigma = 0$ and $\sigma = 1$ hyperplanes.

A curve thus formed as a consequence of Sard's theorem may not simply wander without doing any of the above since it would eventually violate the uniqueness of the curve guaranteed by the implicit function theorem.

Most obviously, simply guaranteeing the existence of smooth, disjoint, non-bifurcating curves does not constitute a method. The Chow-Yorke algorithm is to track the zero curves of $\rho_a(\sigma, z)$ emanating from the (known) zeroes of $g_a(z)$ at the $\sigma = 0$ hyperplane. Thus possibilities iii), iv), and vi) are automatically eliminated.

The procedure then is:

i. Construct a homotopy map $\rho(a, \sigma, z)$ such that the Jacobian matrix $D\rho$ has full rank on $\rho^{-1}(0)$, $\rho(a, 0, z) = g_a(z)$, and $\rho(a, 1, z) = f(z)$. This is where the Chow-Yorke algorithm differs from other homotopic continuation algorithms, for we construct our homotopy not with just one trivial problem, but with a whole range of trivial problems depending on the parameter vector a.

ii. Choose $g_a(z)$ with a unique known root, corresponding to a randomly chosen parameter vector a. Show that the zero curves of ρ_a emanating from $\sigma = 0$ are monotone in σ if $\rho_a(0, z) = g_a(z) = 0$ has more than one solution. Doing so we are able to discard the possibility ii).

iii. Show that the zero curves of $\rho_a(\sigma, z)$ are bounded. This requirement obviates the possibility v), leaving the only remaining choice that the curve will have to cross the $\sigma = 1$ hyperplane.

Thus, the supporting theory[8, 52] guarantees that for almost all a there exists a zero curve from $\sigma = 0$ to a root $\bar{z}$ of $f(z) = 0$ at $\sigma = 1$ and that this curve has finite arc length if $Df(\bar{z})$ is non-singular. The phrase 'almost all a' has the connotation that the set of all those parameter values a for which a homotopy curve fails to exist has Lebesgue measure zero (Sard's theorem). Thus if, perchance, one had the ill luck of choosing one such parameter vector, then an adjacent set of values of a will provide the required curve.

Usually item iii) from the aforementioned procedure is somewhat difficult to show and in some cases may not even be true. So in general a curve starting from $\sigma = 0$ can either

i) reach a root $\bar{z}$ of $f(z)$ at $\sigma = 1$, or

ii) wander off to infinity.

The principal idea behind the Chow-Yorke algorithm is to convert the optimization problem into a system of non-linear equations, the roots of which would yield the solution to the optimization problem. Thus in order to solve a parameter optimization problem, we would have to solve a system of non-linear equations which, e.g., could represent the Kuhn-Tucker necessary conditions.

Suppose we are interested in obtaining a family of solutions with respect to some system parameter. Then in order to use the Chow-Yorke algorithm, we would have to construct a homotopy map with a zero curve connecting the known solution corresponding to a value of the parameter with the solution corresponding to the desired value of the parameter. If we consider the equality expressions and only those inequality constraints which are active, then such a map may not be possible if the inequality constraints which are active change along the homotopy zero curve. Thus, it becomes necessary to consider either some active constraint strategy or somehow include the conditions on the inequality constraints in the form of an equation. We prefer to do the latter.

A way that seems natural is to consider slack variables to convert the inequality constraints into equality constraints and treat them as such. Doing so yields the following sets of equations to be considered along with equations [2.3.3] and [2.3.4]:

$$g_i(\bar{x}, \varepsilon) - \eta_i^2 = 0, \qquad [3.2.1]$$

$$\eta_i \mu_i = 0, \qquad [3.2.2]$$

where $i = 1, \ldots, ni$ and η_i are the slack variables.

However, if the active constraint set changes, we would still obtain erroneous results since we have not included the non-negativity restriction on the inequality constraint multipliers. This last situation may be remedied by using μ_i^2 as the multipliers for the inequality constraints. Thus, by introducing ni additional variables, we have considered the entire set of necessary conditions. On applying this method to our simple example [2.4], we find that the homotopy curve goes off to infinity. Evidently, we are still in need of a better representation of the necessary conditions on the inequality constraints.

Note that the Kuhn-Tucker necessary conditions on the inequality constraints are in the complementarity form. Therefore, we are able to make use of Mangasarian's complementarity theorem.[33]

Mangasarian's Complementarity Theorem[33]

Let $\Theta : \mathcal{R} \rightarrow \mathcal{R}$ *be any strictly increasing function with* $\Theta(0) = 0$.
Then $z \in \mathcal{R}^n$ *and* $f \in \mathcal{R}^n$ *solve the complementarity conditions*

$$z \geq 0, \quad f \geq 0 \quad \text{and} \quad z^T f = 0$$

if and only if z *and* f *satisfy*

$$\Theta(|\, f_i - z_i \,|) - \Theta(f_i) - \Theta(z_i) = 0 \qquad [3.2.3]$$

for $i = 1, \ldots, n$.

One can easily verify the above theorem to be true. Although intended for the situation where $f = f(z)$, it holds true even for two quantities which may not be explicitly

dependent on each other. The simplest function $\Theta(t)$ that yields a C^2 complementarity function is t^3, since $|t|^3$ is C^2.

Thus we can write the Kuhn-Tucker complementarity conditions as:

$$G(x, \mu) \equiv \quad |g_i(x) - \mu_i|^3 - g_i^3(x) - \mu_i^3 = 0, \quad i = 1, \ldots, \text{ni}. \qquad [3.2.4]$$

The above equation along with equations [2.3.3] and [2.3.4] provide the Kuhn-Tucker necessary conditions in the form of equality expressions.

3.3 Homotopy Maps.

Construction of a homotopy map is not straightforward and by no means universal. What works well for one problem may be disastrous with respect to another. However, before attempting to construct fancy homotopy maps, one must try simple maps and if necessary adaptively increase the complexity of the maps to suit the problem.

In solving a parameter optimization problem [2.3.2], dependent on some system parameter ε, one can construct a very simple homotopy map as:

$$\sigma \nabla \mathcal{L}(x, \lambda, \mu, \varepsilon_1) + (1 - \sigma) \nabla \mathcal{L}(x, \lambda, \mu, \varepsilon_0) = 0, \qquad [3.3.1]$$

$$\sigma h(x, \varepsilon_1) + (1 - \sigma) h(x, \varepsilon_0) = 0, \qquad [3.3.2]$$

$$\sigma G(x, \mu, \varepsilon_1) + (1 - \sigma) G(x, \mu, \varepsilon_0) = 0, \qquad [3.3.3]$$

where

$$\mathscr{L}(x, \lambda, \mu, \varepsilon) = C(x, \varepsilon) - \lambda^T h(x, \varepsilon) - \mu^T g(x, \varepsilon),$$

and at $\sigma = 0$, $x = x_0$, $\lambda = \lambda_0$, $\mu = \mu_0$ provide the solution at $\varepsilon = \varepsilon_0$ for the problem [2.3.2].

The number of problems where such a simple homotopy map provides the desired solution is very small. Ideally one imbeds the parameter σ so that the trivial problem would naturally evolve into the desired problem as σ goes from 0 to 1. In other words, the homotopy parameter σ may be imbedded in the system constants, which determine or characterize the system. For instance in the orbital rendezvous problems, the homotopy parameter may be imbedded in the initial conditions of the interceptor and the target so that as σ varies from 0 to 1, the initial conditions obtained are valid starting conditions for some problem. Likewise, the homotopy parameter may be used to define a range of values for the total time required for a maneuver. The idea being that if the solution corresponding to one set of parameters are known, then solution to some other set of system constants are obtained. We shall discuss the space-flight rendezvous problem in greater detail in the succeeding chapters.

Assuming that we are successful in imbedding the homotopy parameter σ intrinsically, we may solve the parameter optimization problem by finding the roots of the Kuhn Tucker equations:

$$\nabla C(x, \sigma) - \lambda^T \nabla h(x, \sigma) - \mu^T \nabla g(x, \sigma) = 0, \qquad [3.3.4]$$

$$h(x, \sigma) = 0, \qquad [3.3.5]$$

$$G(x, \mu, \sigma) \equiv |g_i(x, \sigma) - \mu_i|^3 - g_i^3(x, \sigma) - \mu_i^3 = 0, \quad i = 1, \ldots, \text{ni}. \qquad [3.3.6]$$

$\sigma = 0$ corresponds to the trivial problem and at $\sigma = 1$ we have the desired problem. Note that such a formulation allows one to obtain a family of solutions corresponding to some system parameter of interest.

However, on using such a homotopy map and starting with the solution to a known problem at $\sigma = 0$, a singular Jacobian matrix is encountered at that value of $\sigma = \tilde{\sigma}$ where some constraint changes its active or inactive status. On examining the complementarity function, we note that indeed it produces a zero Jacobian matrix row at the point where an inequality constraint $g_i(x) = 0$ and the corresponding Lagrange multiplier $\mu_i = 0$. Consequently, we have a zero row corresponding to this constraint in the Jacobian matrix of our homotopy map to solve the Kuhn-Tucker equations. Thus, we are once again beset with the old problem of a limited range over which this method can be successfully applied.

Several attempts were made to obviate the above mentioned singularity. This involved using variations on the ideas due to Watson, Bixler, and Poore[56] and some others. The most success was obtained by adding a linear term vanishing at $\sigma = 1$ to the complementarity equations based on an idea due to Watson,[51] i.e.,

$$\sigma \; \alpha \; [|g(x) - \mu|^3 - g^3(x) - \mu^3] \; + \; (1 - \sigma) \; (y - a) = 0, \qquad [3.3.7]$$

where $\alpha = \pm 1$ depending on the constraint and a is chosen initially so that the above equation is satisfied and y is either μ or g. On most of our attempts, $y = \mu$ performed

better than $y = g$. Likewise $\alpha = -1$ for all of the inequality constraints was the most successful.

Thus the homotopy map to solve the parameter optimization problem is given by:

$$\rho_a \equiv \{\nabla \mathcal{L}(x, \lambda, \mu, \sigma),\ h(x, \sigma),\ \sigma[-G(x, \mu, \sigma)] + (1-\sigma)(\mu - a)\}. \qquad [3.3.8]$$

A related homotopy map may be constructed by adding the linear term to the first two terms of the above map giving us the following sets of equations:

$$\sigma\,[\ \nabla C(x, \sigma) - \lambda\ \nabla h(x, \sigma) - \mu\ \nabla g(x, \sigma)\ ] + (1-\sigma)(x - \bar{b}) = 0, \qquad [3.3.9]$$

$$\sigma\,[\ h(x, \sigma)\ ] + (1-\sigma)(\lambda - \bar{c}) = 0 \qquad [3.3.10]$$

$$\sigma\,[\ -G(x, \mu, \sigma)\ ] + (1-\sigma)(\mu - \bar{a}) = 0 \qquad [3.3.11]$$

where $\bar{b} \in \mathcal{R}^n$, $\bar{c} \in \mathcal{R}^{ne}$ and $\bar{a} \in \mathcal{R}^{ni}$.

3.4 *Example.*

Let us reconsider the simple example [2.4]. The homotopy equations [3.3.9], [3.3.10], [3.3.11] for the problem are given by:

$$\begin{aligned} &\sigma\,[2x_1 + x_2 + \mu_1 - 2\mu_2 - 2\mu_3 x_1 + 2\mu_4 x_1] + \\ &(1-\sigma)(x_1 - a_1) = 0, \end{aligned} \qquad [3.4.1]$$

$$\sigma [x_1 + 2[\alpha - 1] x_2 - 2\mu_1 + \mu_2 - 2\mu_3 x_2 + 2\mu_4 x_2] + (1 - \sigma)(x_2 - a_2) = 0, \quad [3.4.2]$$

$$-\sigma [| 2x_2 - x_1 - \mu_1 |^3 - (2x_2 - x_1)^3 - \mu_1^3] + (1 - \sigma)(\mu_1 - a_3) = 0, \quad [3.4.3]$$

$$-\sigma [| 2x_1 - x_2 - \mu_2 |^3 - (2x_1 - x_2)^3 - \mu_2^3] + (1 - \sigma)(\mu_2 - a_4) = 0, \quad [3.4.4]$$

$$-\sigma [| x_1^2 + x_2^2 - 1 - \mu_3 |^3 - (x_1^2 + x_2^2 - 1)^3 - \mu_3^3] + (1 - \sigma)(\mu_3 - a_5) = 0, \quad [3.4.5]$$

$$-\sigma [| 2 - x_1^2 - x_2^2 - \mu_4 |^3 - (2 - x_1^2 - x_2^2)^3 - \mu_4^3] + (1 - \sigma)(\mu_4 - a_6) = 0, \quad [3.4.6]$$

where $\alpha = \sigma\alpha_1 + (1 - \sigma)\alpha_1$.

We start at $\sigma = 0$ with the solution corresponding to $\alpha_0 = 0$ to obtain the solution corresponding to $\alpha_1 = 3$. We used the normal flow algorithm due to Watson, Billups and Morgan,[55] to follow the homotopy zero curve. The path of solutions as σ varies from 0 to 1 is given in Figure 3., superimposed on the figure showing the feasible region. Note that the homotopy zero curve did not satisfy the constraints at all times. The specific variation of the variables $\{x, \mu\}$ with respect to σ is given in Figure 4. The curve tracking algorithm required 69 Jacobian evaluations, and the corresponding arc length of the homotopy zero curve was 3.2978.

We thus have described a way of solving parameter optimization problems which does not involve any active constraint set logic. In the succeeding chapters we shall discuss

the specific application of this Chow-Yorke algorithm based method to the space-flight rendezvous problems.

4.0 Space-flight Rendezvous Problem.

4.1 Equations of Motion

Consider the motion of a satellite about a planet. For the sake of simplicity let us assume that it is isolated from the gravitational influence of other bodies. The equation of motion of this satellite under gravitational influence of the planet is given by:

$$\ddot{\bar{r}} + \frac{\mu}{r^3}\,\bar{r} = \bar{P}, \qquad [4.1.1]$$

where:

- $\bar{r}$ position vector of the satellite at any given time,
- μ the gravitational constant for the planet,
- $\bar{P}$ some external disturbing force per unit mass.

If there were no external disturbances such as air drag and radiation pressure and if we assume that the planet in question is spherical, then in the absence of any thrust the

motion of the satellite is described by equation [4.1.1] with $\bar{P} = 0$. The resulting trajectory is given by a conic section with the focus at the center of the planet[2].

A certain elegance in the equations governing the motion of a satellite about a planet is obtained if one introduces a particular regularizing transformation. In addition to removing singularities that exist in some classical formulations, such a transformation also allows the definition of new orbital parameters, which are related to the constants of motion which describe the orbit in an unperturbed field of influence. These new parameters are desirable since the classical orbital parameters; the semi-major axis a, the eccentricity e, inclination i, longitude of the ascending node Ω, argument of the perigee ω, and the time of perigee passage τ can exhibit singularities in certain circumstances. For instance in an equatorial orbit the longitude of the ascending node is undefined; for a circular orbit the argument of perigee ω is undefined, and in a parabolic orbit the semimajor axis is infinity.

One regularizing transformation which avoids these problems includes replacing the independent variable time with the change in true anomaly η and using the reciprocal of the magnitude of the position vector u and the unit vector $\hat{r}$ in the direction of the position vector instead of the position vector $\bar{r}$ itself.[5,48] Thus the position vector can be represented as:

$$\bar{r} = \frac{\hat{r}}{u}. \qquad [4.1.2]$$

Introducing the definition of the angular momentum h we have

$$\dot{\eta} = h u^2. \qquad [4.1.3]$$

It can be shown that the velocity of the satellite is given by:

$$\overline{V} = \dot{\overline{r}} = h\,[\,u\,\hat{r}' - u'\,\hat{r}\,], \qquad [4.1.4]$$

where $\hat{r}'$ is the derivative of the unit vector $\hat{r}$ with respect to the new independent variable η. If we consider the unit vector along the angular momentum vector,

$$\hat{h} = \frac{\overline{r} \times \overline{V}}{h}, \qquad [4.1.5]$$

then we can show that $\hat{r}$, $\hat{r}'$ and $\hat{h}$ form a set of orthonormal vectors in $\mathscr{R}^3$. By substituting equations [4.1.2] - [4.1.5] into [4.1.1], we can write the following equations for motion for a body moving in an inverse square gravitational field with the perturbation force per unit mass $\overline{P}$:

$$h' = \frac{\overline{P} \cdot \hat{r}'}{h\,u^3}, \qquad [4.1.6]$$

$$\hat{h}' = -\frac{\overline{P} \cdot \hat{h}}{h^2\,u^3}\,\hat{r}', \qquad [4.1.7]$$

$$u'' + u = \frac{\mu}{h^2} - \frac{1}{h^2\,u^3}\,[(\overline{P} \cdot \hat{r})\,u + (\overline{P} \cdot \hat{r}')\,u'\,], \qquad [4.1.8]$$

and

$$\hat{r}'' + \hat{r} = \left(\frac{\overline{P} \cdot \hat{h}}{h^2\,u^3}\right)\hat{h}. \qquad [4.1.9]$$

Note that in the absence of any disturbing force $\overline{P}$, h and $\hat{h}$ are constants and equations [4.1.8] and [4.1.9] reduce to the equations for a simple harmonic oscillator.

Thus for the price of an additional variable, we have a simple set of equations governing the motion of an undisturbed satellite.

Given the initial position, $\bar{r}_0$, and velocity, $\bar{V}_0$, of an orbit, the aforementioned regularized states of the satellite are easily obtained as follows:

i. $u_0 = \dfrac{1}{\|\bar{r}_0\|}$,

ii. $\hat{r}_0 = u_0\, \bar{r}_0$,

iii. $\bar{h}_0 = \bar{r}_0 \times \bar{V}_0$,

iv. $h_0 = \|\bar{h}_0\|$,

v. $\hat{h}_0 = \dfrac{\bar{h}_0}{h_0}$,

vi. $\hat{r}'_0 = \hat{h}_0 \times \hat{r}_0$,

vii. $u'_0 = -\dfrac{u_0}{h_0}\ \{\bar{r}_0 \cdot \bar{V}_0\}$.

Therefore, if the initial conditions are known on any arc along with the change in true anomaly, the conditions at any other point along that arc can be obtained from

$$u(\eta) = \frac{\mu}{h^2} + \left[u(0) - \frac{\mu}{h^2} \right] \cos(\eta) + u'(0) \sin(\eta), \qquad [4.1.10]$$

$$u'(\eta) = -\left[u(0) - \frac{\mu}{h^2} \right] \sin(\eta) + u'(0) \cos(\eta), \qquad [4.1.11]$$

and similarly the properties of the unit vectors are obtained from

$$\hat{r}(\eta) = \hat{r}(0)\ \cos(\eta) + \hat{r}'(0)\ \sin(\eta), \qquad [4.1.12]$$

$$\hat{r}'(\eta) = -\hat{r}(0)\ \sin(\eta) + \hat{r}'(0)\ \cos(\eta), \qquad [4.1.13]$$

$$\hat{h}(\eta) = \hat{h}(0). \qquad [4.1.14]$$

4.2 *The Rendezvous Problem*

One of the interesting problems in space mechanics is that of determining a fuel-optimal rendezvous between two satellites. We shall consider the case where only one of the satellites is capable of maneuvering, it being called the interceptor. The other (passive satellite) shall be called the target. The interceptor therefore is allowed to thrust for short intervals which in turn are followed by longer intervals of non-thrusting or coasting arcs.

Usually thrust arcs are very small in comparison with non-thrust arcs. Consequently, one can approximate the rendezvous trajectory in terms of coasting arcs (zero thrust) separated by a change in velocity at the junction of such arcs.[37,44] This velocity change is equivalent to applying an impulsive thrust at that point.

In terms of the regularized variables that we have chosen, an impulse signifies a change in u' and h, with a change in direction for the states $\hat{r}'$ and $\hat{h}$. Thus using these variables, an impulse vector $\{\Delta V_x, \Delta V_y, \Delta V_z\}$ is characterized by $\{\Delta u', \Delta h, \phi\}$, where $\Delta u'$ is the change in u', Δh is the change in h and ϕ is the angle of rotation of $\hat{r}'$ and $\hat{h}$ about $\hat{r}$ at the point of application of the impulse. Therefore, given a set of variables $\{\Delta u', \Delta h, \phi\}$ for each impulse and the coasting angle η for each arc, and given the initial conditions, i.e., $\{u_0, u'_0, h_0, \hat{r}_0, \hat{r}'_0, \hat{h}_0\}$, the entire trajectory is described.

In general a rendezvous can be accomplished by an infinite selection of combinations of coasting arcs and changes in velocity at the impulse. It is only natural to ask if the maneuver can be performed in some optimal manner. Optimality refers to either fuel economy, or with respect to time or some combination of both.

Suppose that the interceptor and the target were initially residing in circular coplanar orbits and that the ratio of outer to inner radii is less than 15.6,[14,28] then a Hohmann transfer is known to be fuel-optimal. A Hohmann transfer is a two impulse transfer. A tangential first impulse sets the satellite on an elliptic orbit whose perigee radius is equal to the inner circular orbit and the apogee radius as the outer circular orbit. The second tangential impulse provides the rendezvous. If the radii ratio is greater than 15.6 then a Hohmann transfer is found to be no longer fuel-optimal. A bi-elliptic transfer,[21] which is a three impulse transfer, can be found yielding a lower cost function value than the Hohmann transfer. Ideally, the first impulse results in a parabolic orbit going to infinity where an infinitesimal impulse is applied which initiates the return to the target orbit, where the third impulse provides the rendezvous. Although, such a transfer provides the absolute minimum for coplanar circle-to-circle rendezvous problems with radii ratio greater than 15.6, any three impulse transfer going first to a radius larger than the outer orbit and then coming back is more fuel efficient than the Hohmann transfer. This scenario is one of the first well known cases where a multiple impulse transfer is better than a simple two impulse transfer.

If there is no restriction on the fuel consumed, then a time optimal rendezvous with impulsive thrust is one with infinite thrust along the shortest line joining the target and the starting point. The problem becomes interesting if an inequality type restriction is placed on the total time to rendezvous and one considers the fuel-optimal problem.

This problem can be cast in the form of a parameter optimization problem of choosing the variables - coast angle η on an arc, and $\{\Delta u', \Delta h, \phi\}$ denoting an impulse, so that a function representing the fuel consumed is minimized. We would require that the

solution so obtained satisfies certain constraints such as the rendezvous conditions, and that the time to rendezvous does not exceed a prescribed maximum time limit.

4.3 Parameter Optimization Problem

The problem of interest is to find a minimum fuel rendezvous trajectory between two bodies (one of which is non-maneuvering) initially residing on Keplerian orbits in terms of impulses (magnitude and direction and their respective firing times). In other words, we are seeking a sequence of coasting (Keplerian) arcs separated by changes in velocity (impulsive thrust), which would provide the rendezvous between the interceptor and the target. It can be shown that the fuel consumed is related directly to the characteristic velocity,[28, 34, 35] which is the sum of the absolute change in velocity at each impulse. The problem then is one of minimizing the characteristic velocity.

At an impulse, the change in velocity vector is given by:

$$|\Delta \overline{V}_i| = \overline{V}_i^+ - \overline{V}_i^- = \overline{V}_{i+1}(0) - \overline{V}_i(\eta), \qquad [4.3.1]$$

where $\overline{V}^+$ is the velocity of the interceptor after an impulse and hence the velocity at the beginning of the next coasting arc, and $\overline{V}^-$ is the velocity of the interceptor before the impulse, i.e., the velocity at the end of the previous coasting arc. Thus in terms of the regularized variables the magnitude of the velocity change vector at the i^{th} impulse is given as

$$|\Delta \overline{V}_i| = \sqrt{u_{i+1}^2(0)\,[h_{i+1}^2 - 2\,h_i\,h_{i+1}\,\cos(\phi_i) + h_i^2] + [\Delta h_i\,u'_{i+1}(0) + \Delta u'_i\,h_i]^2}\,. \qquad [4.3.2]$$

and the characteristic velocity for the complete rendezvous maneuver is given by

$$C = \sum_{i=1}^{\text{nim}} |\Delta \overline{V}_i|, \qquad [4.3.3]$$

where nim is the number of impulses.

The number of impulses has to be prespecified. Hence any solution which minimizes [4.3.3] is only a locally optimal solution for the prespecified number of impulses. Unfortunately, there is no clue indicating whether more impulses would be better. Therefore, the best one can hope to do is keep on increasing the number of impulses and successively solve the parameter optimization problem until there is no discernible change in cost, or until an impulse 'collapses' to zero. In other words, if at the end of the optimization procedure one of the intermediate coast angles goes to zero or the total change in velocity approaches zero, then for that problem one less impulse will suffice for a locally optimal solution. Thus the correct number of impulses required for a problem is known only after an above mentioned situation occurs. Clearly, one would like some way of knowing whether a solution thus obtained is truly optimal other than knowing that it is the best possible scenario for the prescribed number of impulses. To this effect one can invoke conditions for an optimal trajectory using the Maximum Principle applied to this problem and check the conditions against the trajectory obtained.[57,58]

Mathematically, the above parameter optimization problem can be described as choosing a sequence of $\{\eta, \Delta u', \Delta h, \phi\}$ so that the characteristic velocity (total velocity change) is minimized. Thus the time limited problem can be stated as:

$$\underset{S}{\text{Min}} \quad C(x), \qquad [4.3.4]$$

where $S = \{ (\eta, \Delta u', \Delta h, \phi)_j, \eta_t \ , j = 1, \ldots, \text{nim}, \}$, where nim = prespecified number of impulses, and η_t is the angle of coast for the target.

In addition, the following equality and inequality constraints must be satisfied:

Equality Constraints

The conditions for rendezvous require the following position and velocity matching constraints:

i. final position match constraint -

$$h_1(x) \equiv \quad \vec{r}_f \ - \ \vec{r}_t(\eta_t) = 0, \qquad [4.3.5]$$

ii. final velocity match constraint -

$$h_2(x) \equiv \quad \vec{V}_f \ - \ \vec{V}_t(\eta_t) = 0, \qquad [4.3.6]$$

iii. time of flight match constraint -

$$h_3(x) \equiv \quad T_f \ - \ T_t \ = \ 0, \qquad [4.3.7]$$

where the subscript f refers to the conditions on the interceptor trajectory after the final impulse and the subscript t refers to conditions on the target.

Inequality Constraints

Additional constraints which must be avoided along each arc of the interceptor or target trajectory in the form of an inequality are:

i. non-negativity of the coasting arcs of the interceptor -

$$g_i(x) \equiv \quad \eta_i \geq 0 \qquad i = 1, \ldots, \text{nim}, \qquad [4.3.8]$$

ii. non-negativity of the coasting arc of the target -

$$g_{\text{nim}+1}(x) \equiv \quad \eta_t \geq 0, \qquad [4.3.9]$$

iii. time of flight limit constraint (maximum time specified for rendezvous) -

$$g_{\text{nim}+2}(x) \equiv \quad T_{\max} - T_f \geq 0, \qquad [4.3.10]$$

iv. minimum radius constraint for each coasting arc except the initial coast arc of the interceptor trajectory -

$$g_j(x) \equiv \quad u_0 - u_j(x) \geq 0 \quad j = \text{nim} + 3, \ldots, 2\,\text{nim} + 1. \qquad [4.3.11]$$

The transfer arc should lie outside a circle of radius $r_0 \equiv \frac{1}{u_0}$.

v. non-negativity of the radius constraint -

$$g_j(x) \equiv \quad u_j(x) \geq 0 \quad j = 2\,\text{nim} + 2, \ldots, 3\,\text{nim}. \qquad [4.3.12]$$

In the formulation that we have chosen to describe the trajectories, if a large transfer angle is specified on a hyperbolic orbit then it is mathematically possible to obtain a negative value for the reciprocal of the magnitude of the

radius vector u. A negative u is physically meaningless since it would mean a negative radius. Hence the requirement of this constraint.

Constraints iv. and v. are essentially infinite inequality constraints to be satisfied over the whole interval $[0, \eta_j]$ on the j^{th} coasting arc. Consequently, our problem falls under the broad class of semi-infinite optimization.[39] However, following a proposal due to Polak,[39] we recast the above infinite constraints using max-functions. Intuitively, this makes sense. Consider the infinite constraint iv., i.e., requiring the entire coasting arc to lie outside a circle of radius r_0. Thus, if the minimum radius that occurs on a coasting arc satisfies the minimum radius constraint, then obviously the entire coasting arc would satisfy the constraint. Thus, all that is required is to monitor the minimum radius occurring on the coasting arc, and require that it satisfy the constraint iv.

From the nature of the transfer arcs, i.e., conic sections, the minimum radius on any subarc is given by:

$$\frac{1}{u_{max}} = \begin{cases} \text{perigee radius, if perigee passage occurs on subarc,} \\ \min\ (r_{initial}, r_{final}), \text{ otherwise} \end{cases}$$

The minimum radius constraint as given above is not C^2 and consequently for the moment we have chosen the stiffer constraint of requiring the perigee radius of any transfer arc to be greater than than the minimum allowable radius. This has certain pitfalls which will be discussed later.

Likewise, the non-negativity of the radius constraint v. must also be recast in a more tangible form. One way to reformulate the above constraint is to ensure that the minimum value of $u_i(\theta)$, $\theta \in [0, \eta_i]$, is positive. We prefer to ensure that the final value of $u_i(\theta)$ on the i^{th} coasting arc be non-negative, with the assumption that transfer

angles greater than twice the asymptote angle $\eta_{asymptote} = \cos^{-1}(\frac{-1}{e})$ would not be encountered if some of the coasting arcs are hyperbolic.

4.4 *Minimum Radius Constraint*

As stated above, the minimum radius constraint is of the infinite type. One of the ways the infinite constraints are handled is with the use of max-functions. Thus for our problem the minimum radius constraint can be written as:

$$U_0 - u_{max} \geq 0. \qquad [4.4.1]$$

For non-circular subarcs, if $0 \leq \eta_{perigee} \leq \eta$, then

$$u_{max} = U_{perigee} = \frac{\mu}{h^2} + \sqrt{\left[u(0) - \frac{\mu}{h^2}\right]^2 + u'^2(0)}, \qquad [4.4.2]$$

or else

$$u_{max} = \max \{ u(0), u(\eta) \}. \qquad [4.4.3]$$

For a circular orbit $u_{max} = u(0) = u(\eta) = \frac{\mu}{h^2}$.

As is self evident, $u_{max}(x)$ is not C^2. It is continuous but does not have continuous derivatives. For most cases, the transfer subarcs are such that the perigee radius is greater than or equal to the minimum allowable radius. However, there are some interesting trajectories, which involve an outward impulse to set the interceptor on a

near elliptic rectilinear orbit. The intercept occurs on its downward journey where the target orbit intersects the rectilinear orbit. A second normal impulse provides the rendezvous. As one can guess, these trajectories are very fuel-inefficient, but these do indeed satisfy the first order necessary conditions.[57,58]

Discounting such trajectories, we require the perigee radius on each of the component subarcs (except the wait on the initial parking orbit) to satisfy the minimum radius constraint. Thus the minimum radius constraint becomes:

$$U_0 - \frac{\mu}{h^2} - \sqrt{\left[u(0) - \frac{\mu}{h^2}\right]^2 + u'^2(0)} \geq 0. \qquad [4.4.4]$$

The minimum radius expression is C^2 except on circular orbits, i.e., when $u(0) = \frac{\mu}{h^2}$ and $u'(0) = 0$, where the max-function is not defined.

Let us now endeavor to modify this expression so that it retains its original characteristics when not on a circular orbit, but yields the finite derivatives for a circular arc. Such a modification would allow us to consider a wide range of problems with the use of only one expression. Most of our attempts involved either adding a term inside the square root term which would remove the singularity that exists on the circular arcs, or multiplying the square root term with some function that essentially provides a C^2 function over the whole domain. The first consideration straight away obviates certain interesting trajectories that in actuality include constrained arcs, whereas the second type of functions under-estimated the max-function. Thus, it is possible for the trajectory to violate the true minimum radius constraint even though it does not violate the stated constraint. The following function provides the best function to date which has the properties we are seeking, i.e., is C^2 over the whole domain of circular and non-circular

orbits, and gives a good approximation to the original constraints for both circular and non-circular orbits.

$$\frac{\tanh^2(\gamma\, e^2)}{U_{\text{perigee}}} + \frac{h^2}{\mu}\,[1 - \tanh^2(\gamma\, e^2)] - \frac{1}{U_0} \geq 0 \qquad [4.4.5]$$

where U_{perigee} is given by [4.4.2], and e is the eccentricity given by:

$$e = \frac{h^2}{\mu}\sqrt{\left[u(0) - \frac{\mu}{h^2}\right]^2 + u'^2(0)}\,. \qquad [4.4.6]$$

Although this expression does indeed under-estimate the minimum radius constraint, the true constraint could be approximated as closely as required with the introduction of another homotopy in the expression such as $\gamma = \sigma\,\gamma_1 + (1-\sigma)\,\gamma_0$. We have had reasonable success with this formulation. In succeeding chapters we shall construct the specific homotopy map required to solve a variety of orbital rendezvous problems.

5.0 Applications to Space-flight Rendezvous Problems

5.1 Introduction

Having discussed the numerical algorithm in the preceding chapters, we now proceed to apply it to the space-flight rendezvous problem. A homotopy approach is especially useful since it facilitates the study of the effect on the fuel consumption or equivalently the characteristic velocity due to variations in parameters such as the time of flight limit, minimum radius limit, the inclination of the target orbit with respect to the interceptor orbit, phase difference between the target and the interceptor, etc.

Ideally we would select some homotopy map which would take us from the solution for a simple problem, which has few constraints and can be determined analytically, to the solution for a complex problem which may have several constraints and may include out-of-plane maneuvers. Furthermore it would be desirable to have every solution along

the homotopy path satisfy the necessary conditions for an optimal maneuver for some rendezvous problem. One of the main reasons for embarking on the homotopy approach to solve parameter optimization problems is in an attempt to dispense with the need for a good initial guess, or at least seek a way of enlarging the domain of convergence for an initial guess.

The homotopy map given by equations [3.3.4]-[3.3.6] has this desirable property and all solutions along the homotopy path are solutions to legitimate rendezvous scenarios. For discussion purposes we will designate this map as map 1. However, as described earlier the Jacobian matrix associated with this map becomes singular when a trajectory initially contacts a constraint and both the constraint and the associated multiplier are zero at the same time. Until such a situation arises, the sequence of solutions obtained are valid. Hence whenever the active constraint set changes, an alternative procedure is required and a different but related homotopy map is used which allows the active constraint set to change and has the property that it converges to a legitimate problem only at the final value of the homotopy parameter, the intermediate results giving no useful information. This map is given by equations [3.3.9]-[3.3.11] and will be designated map 2. It is convenient then to use the two maps to undertake this study, the first to generate a family of solutions which have the same active constraint set, and the second to bridge the point where the constraint set changes.

For example, one might start with a simple Hohmann transfer which has a known analytical solution that consists of a two impulse transfer between two coplanar circular orbits. The solution to this problem can then be used as an initial solution for a homotopy to a problem which has a restricted time for rendezvous. The maximum time of flight can be included in the homotopy map as will be described in the next section.

If map 1 is used, every solution along the zero curve is a solution to a rendezvous problem with a specified maximum time of flight. As the maximum time of flight is reduced, generally the minimum radius constraint will become active. If this occurs, the Jacobian matrix associated with the first map will become singular. At this point any previous solution along the zero curve may be used as an initial solution in map 2 to obtain the solution across the singularity. A slightly shorter time is selected and the homotopy path tracked to its completion ($\sigma = 1$). The result in general will include a solution to a rendezvous problem with the minimum constraint active. Once the solution has been obtained, it can serve as an initial solution for the original homotopy, map 1, as the maximum time for the maneuver is further shortened.

A similar procedure can be used for variations in phase angle, plane inclination, or some other parameter. The results, will give two impulse trajectories which satisfy the Kuhn-Tucker necessary conditions. If a three (or more) impulse trajectory is required, the initial problem must be made to look like a three (or more) impulse problem. It has been found that this requirement is best achieved by splitting the single coast arc of the Hohmann (or any other solution) into two or more equal parts separated by a zero impulse. This procedure works significantly better than splitting an impulse into two impulses separated by an arc of length zero. These changes may alter the character of the problem, for example, the active constraint set may be different from that of the trivial problem. It is advisable to start with map 2 for this reason.

Details for constructing these maps by imbedding a set of parameters intrinsically into the system constants is presented in the next section.

5.2 *Homotopy Map for the Space-flight Rendezvous Problem.*

Let us now apply the Chow-Yorke algorithm to the space-flight rendezvous problem. The initial conditions of the interceptor at the start of a maneuver are described by $u(0)$, $u'(0)$, $h(0)$, $\hat{r}(0)$, $\hat{r}'(0)$, and $\hat{h}(0)$. If the changes in true anomaly for each subarc are known along with the changes $\Delta u'$ and Δh in u and h along with the out of plane rotation ϕ of $\hat{r}'$ and $\hat{h}$ at the intersection of each coasting arc, then the entire trajectory can be computed easily using the equations [4.1.10]-[4.1.14].

Suppose that we are interested in finding a family of solutions to the minimum fuel spaceflight rendezvous problem for different initial conditions of the interceptor. Then the natural way to imbed the homotopy parameter σ would be a linear map for the scalar initial conditions, i.e.,

$$u(0) = \sigma u_1(0) + (1-\sigma) u_0(0), \qquad [5.2.1]$$

$$u'(0) = \sigma u'_1(0) + (1-\sigma) u'_0(0), \qquad [5.2.2]$$

$$h(0) = \sigma h_1(0) + (1-\sigma) h_0(0). \qquad [5.2.3]$$

where $(\,)_0$ designate the initial conditions for the known problem and $(\,)_1$ denotes the initial conditions for the desired problem.

An additional homotopy map is needed for the unit vector triad, but a linear homotopy map is not advocated since it will not preserve the norm (property of the unit vector).

To create a map, we note that the unit vector triad $\{\hat{r}(0), \hat{r}'(0), \hat{h}(0)\}$ for any problem can be obtained by a sequence of simple Euler angle rotations of (χ, ζ, ψ). Here we will use the 3, 2, 1 sequence to represent the rotation from $\{\hat{r}_0(0), \hat{r}'_0(0), \hat{h}_0(0)\}$ to $\{\hat{r}_1(0), \hat{r}'_1(0), \hat{h}_1(0)\}$. Thus

$$\begin{bmatrix} \hat{r}(0) \\ \hat{r}'(0) \\ \hat{h}(0) \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & \cos(\sigma\chi)\,I & \sin(\sigma\chi)\,I \\ 0 & -\sin(\sigma\chi)\,I & \cos(\sigma\chi)\,I \end{bmatrix} \begin{bmatrix} \cos(\sigma\zeta)\,I & 0 & -\sin(\sigma\zeta)\,I \\ 0 & I & 0 \\ \sin(\sigma\zeta)\,I & 0 & \cos(\sigma\zeta)\,I \end{bmatrix}$$
$$\begin{bmatrix} \cos(\sigma\psi)\,I & \sin(\sigma\psi)\,I & 0 \\ -\sin(\sigma\psi)\,I & \cos(\sigma\psi)\,I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} \hat{r}_0(0) \\ \hat{r}'_0(0) \\ \hat{h}_0(0) \end{bmatrix} \qquad [5.2.4]$$

where ψ is the rotation about the $\hat{h}_0(0)$ axis, ζ is the rotation about the $\hat{r}'$ axis and χ is the rotation about $\hat{r}$ axis to obtain $\{\hat{r}_1(0), \hat{r}'_1(0), \hat{h}_1(0)\}$ of the desired problem. The angle ψ is given by

$$\psi = \tan^{-1}\left[\frac{\hat{r}_1(0) \cdot \hat{r}'_0(0)}{\hat{r}_1(0) \cdot \hat{r}_0(0)}\right]. \qquad [5.2.5]$$

ζ is obtained from

$$\sin(\zeta) = -\hat{h}_0(0) \cdot \hat{r}_1(0), \qquad [5.2.6]$$

with

$$\cos(\zeta) = \frac{\hat{r}_0(0) \cdot \hat{r}_1(0)}{\cos(\psi)}, \quad \text{or} \quad \cos(\zeta) = \frac{\hat{r}'_0(0) \cdot \hat{r}_1(0)}{\sin(\psi)}, \qquad [5.2.7]$$

depending on the value of ψ. χ is determined by

$$\chi = \tan^{-1}\left[\frac{\hat{h}_0(0) \cdot \hat{r}'_1(0)}{\hat{h}_0(0) \cdot \hat{h}_1(0)} \right]. \qquad [5.2.8]$$

In a like manner, we introduce the homotopy parameter σ in the initial conditions of the target. Since we are also interested in obtaining a family of solutions corresponding to different prescribed maximum time to rendezvous limits, a linear homotopy can be introduced as follows:

$$\sigma T_{\max 1} + (1 - \sigma) T_{\max 0} - T_I \geq 0, \qquad [5.2.9]$$

where

T_I is the total time to rendezvous for the interceptor,

$T_{\max}$ is the prescribed maximum total time to rendezvous limit.

A similar linear homotopy map is also introduced for the minimum radius constraint for each transfer subarc,

$$\sigma U_1 + (1 - \sigma) U_0 - u_{\max} \geq 0, \qquad [5.2.10]$$

where

$\frac{1}{U}$ is the prescribed minimum allowable radius limit.

$u_{\max}$ is given by equation [4.4.5].

The subscript 1 refers to the conditions for the final problem and the subscript 0 refers to the conditions for the trivial case.

Note that as σ varies from 0 to 1, equations [5.2.1]-[5.2.8] yield feasible initial conditions for some maneuver, and [5.2.9] and [5.2.10] provide valid constraint limits.

The constraint limits such as U and T_{max}, the Euler angles (χ, ζ, ψ), the initial conditions $\{u_0, u'_0, h_0\}$ of the trivial problem and the desired problem, and $\{\hat{r}_0(0), \hat{r'}_0(0), \hat{h}_0(0)\}$ of the trivial problem serve as the parameters that naturally deform the trivial problem into the desired problem. Using the above imbedding one can use the first homotopy map, map 1, given by equations [3.3.4]-[3.3.6] to obtain solutions corresponding to variations in one or more of the above mentioned system constants. If a singular Jacobian matrix is encountered, which may occur at a point where some inequality constraint changes its active or inactive status, then we may use the second homotopy map, map 2, given by the equations [3.3.9]-[3.3.11]. One should note that there is no guarantee that the solution to the map 2 is in the same family, i.e, a continuation of the solutions found up to the point where the active constraint set changed. Similarly, the last mentioned homotopy map may also be used if one is simply interested in obtaining a solution belonging, perhaps, to a completely different family of solutions. In the use of these maps we do, however, hope that the Jacobian matrix of the second homotopy map is non-singular at $\sigma = 1$, which corresponds to the Kuhn-Tucker equations. In other words, we hope that we are not unlucky in picking the desired problem to be one for which some constraint and its associated Lagrange multiplier are both zero at the final value of the homotopy parameter σ.

The homotopy equations to be solved, either map 1 or map 2, require the gradients of the cost function, equation [4.3.3], and the constraints, equations [4.3.5]-[4.3.12]. The above mentioned gradients and their derivatives with respect to the variables $\{[\eta, \Delta u', \Delta h, \phi]_i, \eta_T\}$, $i = 1, \ldots,$ nim are obtained using the chain rule. The time of flight derivatives and the derivatives of the initial conditions on any subarc are provided in the appendix. Thus derivatives of the minimum radius constraint on any arc given by equation [4.4.5] may be obtained using the chain rule by first taking the derivatives

with respect to the initial conditions. In a like manner first and second derivatives for the other constraints may also be obtained. The curve tracking was performed using the normal flow algorithm FIXPNF from the set of algorithms given in HOMPACK due to Watson, Billups and Morgan.[55] The algorithm requires the computation of the Jacobian matrix of the homotopy map to be tracked.

5.3 Illustration of the use of map 2.

One of our objectives was to examine if we could obtain multi-impulse transfers by starting from a Hohmann transfer solution. The idea then is to

1) split any of the coasting arcs and introduce a zero impulse between them, or

2) split any impulse and introduce a zero coast angle.

Let the interceptor initially reside on a circular orbit of radius 1.2 DU (where 1 DU = 6378.145 km), and the target on a circular orbit of radius 1.45 DU. Let the initial phase angle of the target with respect to the interceptor be 90°. For the Hohmann transfer to occur, the interceptor waits in its own orbit until the relative phase angle of the target becomes 0.3974 radians (22.7668°) at which time the first impulse (purely tangential) is applied. The intercept occurs after the interceptor has moved through 180° at which time the second (also tangential) impulse provides the rendezvous.

For the above example, if the initial wait angle of 4.7482 radians (272.0554°) is split with a zero impulse separating them, then the solution to the desired three impulse problem corresponding to a time to rendezvous limit = 10.411 TU (140 minutes) is obtained by tracking the homotopy equations [3.3.9]-[3.3.11]. The curve tracking procedure required 86 Jacobian function evaluations, and the resulting homotopy curve has an arc length = 3.981. The resulting three impulse trajectory comprises of an initial wait of 0.3351 radians (19.1998°), an inward impulse of 0.2263×10^{-2} DU/TU, a coast of 4.2427 radians (243.0888°) , a second impulse of 0.4315×10^{-1} DU/TU onto the intercept trajectory, a coast of 2.9559 radians (169.3606°), and the final impulse of 0.4368×10^{-1} providing the rendezvous.

On the other hand, if one of the intermediate impulses or the smaller second transfer arc was split, we were unsuccessful in obtaining a solution. Usually, when an intermediate impulse or a coasting arc is split, the homotopy path either turns back and crosses the $\sigma = 0$ hyperplane or yields a degenerate solution, i.e., the final solution has the same characteristics as the trivial problem supplied (i.e., zero coast or zero impulse components).

This procedure was successful in providing three impulse trajectories from some other two impulse family of solutions, which involve an initial coast before the first impulse.

5.4 *Circle-to-Circle Rendezvous Examples.*

5.4.1 Coplanar Case.

Whereas several theoretical and numerical multi-impulse orbital rendezvous trajectories[14] have been described for the unlimited total time case, only a few such solutions[6,7] are available numerically or otherwise if a restriction is placed on total time for the maneuver.

If the initial orbits of the interceptor and the target are circular, then the parameters that characterize the transfer arc are the initial interceptor and target orbit radii, phase difference at the start of the maneuver β_T (i.e., the difference between the true anomaly of the target and the interceptor), the time to rendezvous limit T_{max}, the minimum radius limit r_0, and the inclination of the target orbit with respect to the interceptor orbit i_T. Let us fix the interceptor and target orbit radii at 1.2 DU and 1.45 DU, and the minimum radius limit r_0 = 1.0DU, and successively examine the variation of the cost function value, i.e., the characteristic velocity, with respect to the total time to rendezvous for different target phase angles β_T.

To generate the variation of the cost function with respect to the total time to rendezvous, we use the homotopy map given by equations [3.3.4]-[3.3.6] and imbed the initial and final time to rendezvous limit using equation [5.2.9] with all other system parameters such as the minimum radius limit U, initial conditions of the interceptor and that of the target held fixed. The variation of the characteristic velocity with respect to the total time to rendezvous are given in Figures 5-8 for the relative phase angle of the

target at start $\beta_T = 0°, 90°, 180°$, and 270°. All solutions were obtained in a systematic manner starting from the solutions made available from an earlier effort. These figures agree with those obtained by Chiu[6,7] and those obtained earlier.[57,58]

In general, fuel-optimal rendezvous trajectories are composed of near tangential velocity impulses. Therefore, the interceptor trajectories are mainly characterized by whether the interceptor needs to catch up to the target or 'speed up' or if it has to wait for the target or 'slow down' so that the last two impulses resemble a Hohmann transfer. Consequently for a rendezvous, proper relative phase angle must be attained prior to the penultimate impulse. The former (speed up) type of trajectories (henceforth called type A for easier referral) involve an inward impulse to bring the interceptor within the inner orbit, and the proper relative phase angle is attained while residing on these inner coasting arcs. As may easily be guessed, such trajectories are bounded by the minimum radius constraint, and with a reduced total time to rendezvous limit touch point arcs, (arcs that touch the minimum radius constraint), and constrained arcs, (arcs which follow the minimum radius constraint for a finite time) are encountered. The cost variation of these trajectories are given by the solid lines in each of the Figures 5-8, where * represents the time of fight restriction less than which the minimum radius constraint becomes active.

The latter (slow down) type, (henceforth called type B), involves an immediate outward thrust. These resemble the classical bi-elliptical orbit transfers of Hoelker and Silber.[21] The cost variation for these trajectories are given by the dotted lines in the Figures 5-8.

An example set of trajectories for the unrestricted time problem involving 2, 3, and 4 impulse trajectories of type A are shown in Figures 9-11 and of type B for the 2, and 3 impulse trajectories are shown in Figures 12 and 13, where the diamond indicates the

location of an impulse and the location at start of the maneuver for the interceptor, and an O indicates the location of the target at the start and end of the maneuver. Although we have not found any 4 impulse trajectories of type B, we do not rule out their existence. All the trajectories shown here correspond to the target starting phase angle $\beta_T = 0°$. For this case, as can be seen from the Figure 5, the 3 impulse trajectories were found to be more fuel efficient than the 4 impulse trajectories of type A.

As may be noted in Figures 5-8, the variation of the cost function of the type A orbits with respect to the time to rendezvous for different initial relative phase angle of the target is laterally shifted with respect to the time to rendezvous. This observation is not entirely surprising since the principal part of the maneuver is setting up the proper relative phase angle prior to the penultimate impulse. Therefore, the phasing time may vary while the total cost may not change very much depending on the initial phase angle of the target.

The two impulse family corresponding to the type B transfers was determined by starting with a solution corresponding to the Hohmann transfer. This phase angle β_H is given by

$$\beta_H = \pi\left[1 - \sqrt{\frac{1}{8}\left(1 + \frac{1}{R}\right)^3}\right], \qquad [5.4.1]$$

where R is the ratio of the orbit radii. To study the variation of the characteristic velocity with the initial target phase angle, the time to rendezvous limit T_{max} is held fixed at a large number, and the initial condition homotopy map for the target given by the equations [5.2.1]-[5.2.8] is used. The solutions to the homotopy map, map 1, given by equations [3.3.4]-[3.3.6] provide the desired result. As the phase angle is decreased until the starting relative phase angle is zero and decreased further when the interceptor leads

the target, the time to rendezvous for the unbounded time maneuver first decreases then increases monotonically. Whereas the cost for the maneuver increases monotonically over the same range of variation of the starting relative phase angle β_T as shown in Figure 14. It is interesting to note that there are a number of type B transfers for the same relative initial phase angle of the target, all of which satisfy the Kuhn-Tucker necessary conditions. A similar characteristic is noted for the three impulse case also.

5.4.2 Non-coplanar Case.

Let us next examine the variation of the characteristic velocity with respect to time of flight for different inclinations of the target orbit. The orbital rendezvous problem for the inclined cases present an interesting situation, since the linear analysis due to Neustadt[35] indicates that these trajectories may require more than four impulses for the fuel-optimal time-free maneuver. For the time free case, transfers between inclined hyperbolic orbits was extensively examined by Gobetz[18]. He presents some six impulse transfers for this problem. Most of such rendezvous schemes described in ref 18. require going to infinity, where some infinitesimal impulses are applied to bring the interceptor to the target orbit, where the final impulse provides the rendezvous. However, when the total time to rendezvous is limited by an upper bound, which is indeed the case for problems of more practical nature, then such transfers are relegated to something only of academic interest. Nevertheless, such transfers often hint as to how certain maneuvers may be carried out.

For the cases considered here, at the start of the maneuver the interceptor begins at the line of nodes at the ascending node. The target position is described by the phase angle

which for the non-coplanar cases is defined as the angle between the target position and the line of nodes measured from the ascending node in the direction of the target motion.

Typically, multiple (more than two) impulse rendezvous trajectories that were found locally optimal involved a few small out-of-plane impulses with at least one principal out-of-plane impulse. The different families of rendezvous trajectories differed in the order in which the impulses were applied and also the location of the final impulse. Interestingly, locally time-free fuel-optimal rendezvous trajectories usually involved an application of the impulses at or near the line of nodes (i.e., the line of intersection of the two orbit planes).

The two impulse families provided some interesting trajectories. One family of trajectories involves an initial wait on the interceptor orbit and a small out-of-plane impulse almost tangential to the target orbit at the line of nodes, followed after suitable coast by a second impulse, with a dominant normal component, providing the rendezvous. The second family that was found involves an initial out-of-plane impulse at the line of nodes, a coast, followed by a near tangential impulse providing the rendezvous with the target. The characteristic velocity of the second family of trajectories for the rendezvous depends very much on the position of the target at start. The cost variation with respect to the time to rendezvous are given by the dotted lines in the Figures 15-26.

The variation of the cost function with respect to various target inclination are shown in Figures 15-26 for different starting position of the target. As may be seen the cost of the maneuvers family increases with inclination. This is only natural, since most of the change is required in simply providing the change in inclination. Consider a pure

plane change maneuver. Equation [4.3.2] giving the magnitude of the change in velocity may be rewritten for this special case as

$$|\Delta\overline{V}| = 2\,u\,h\,\sin(\frac{\phi}{2}) = 2\,V_{\text{tangential}}\,\sin(\frac{\phi}{2}), \qquad [5.4.2]$$

where u, h are the conditions at the impulse. Thus for a 60° plane change on a circular orbit of radius 1 DU, an impulse of 1 DU/TU. is required, which is of the same order of magnitude as the cost for some coplanar maneuvers. Note that the cost for a pure plane change maneuver depends on u, i.e., the change in velocity at the impulse is proportional to the reciprocal of the magnitude of the position vector. Therefore, we should expect that any large plane change maneuvers should occur as far from the center of the Earth as allowed by the total time to rendezvous limit. In fact, for an unbounded time maneuver, the fuel optimal rendezvous between two circular orbits of same radius but inclined to each other involves a tangential impulse to an orbit of radius vastly greater than the initial radius, where an impulse provides the plane change and sets the interceptor on a Hohmann transfer to the target orbit. The final small tangential impulse provides the rendezvous.[14]

A typical set of trajectories for the target inclination of $i_T = 0.2$ radians for two, three and four impulse trajectories are shown in Figures 27-32. For this case the starting target phase angle $\beta_T = 270°$. The dotted line shows the Earth and the equatorial plane. The time of flight constraint is inactive for all the trajectories shown for this case. Note that the rendezvous impulse in all but one of the trajectories occurs near the line of nodes. Thus the main difference in each of these trajectories is the time required for the target phase angle to be such that the intercept can occur at the line of nodes.

Figures 33-41 show the different families of rendezvous trajectories for the target inclination of 1.0 radian (57.29°), with the starting target phase angle $\beta_T = 0°$. As with the $i_T = 0.2$ radian case, these trajectories also differ from each other mainly in the time required for the target phase angle to be such that the intercept would occur at the line of nodes. However, since the inclination of the target orbit is large, the location at which the principal out of plane maneuver is applied becomes significant. One three impulse trajectory obtained for this case which resembles the Marchal transfer,[14] provides the minimum for all time to rendezvous greater than 10.7 TU but less than 21.93 TU, and involves a small out of plane maneuver (0.12 × the total cost), coasts for about half revolution, then applies the principal out of plane impulse (0.83 × the total cost) at a radius of 1.62 DU to set it on an intercept trajectory, coasts for about half a revolution for the intercept, where the small final impulse (0.05 × the total cost) provides the rendezvous. This trajectory is shown in Figure 39. If the total time to rendezvous limit is increased then another trajectory (with time of flight constraint inactive) similar to the above trajectory is obtained (Figure 40), where now the second impulse is applied at about a radius of 3.88 DU. This trajectory requires 21.93 TU for the maneuver, but consumes less fuel for the maneuver than the trajectory shown in Figure 39. If the total time to rendezvous limit is increased further, another time unbounded trajectory is obtained which requires even less fuel. A simple computation of the cost incurred to set the interceptor on to a parabolic trajectory to infinity where an infinitesimal impulse provides the necessary out-of-plane maneuver to set it back to the target shows a higher cost than required by the above two trajectories. Consequently, it is suspected that there is a finite total time-to-rendezvous limit which produces the global minimum for unbounded time maneuver of this type, which involves a thrust to a radius of some finite radius, a plane-change thrust on to a near Hohmann trajectory nearly in the plane of the target, and the final impulse producing the rendezvous. The cost corresponding to this

type of three impulse trajectories are shown by the dotted lines in Figures 15-26. This particular type of transfer seems to provide the minimum-fuel trajectories for the case of target inclination $i_T = 1$ radian. However, if the target inclination is reduced to $i_T = 0.5$ radian, these transfers no longer provide the minimum-fuel trajectories for all target starting conditions and total time to rendezvous limits, and it may be better to wait in some intermediate orbit for the target phase angle to be such that the intercept can occur at the line of nodes where the final impulse (with a relatively large out-of-plane component) provides the rendezvous.

5.5 General Orbital Rendezvous Examples.

If the interceptor and target orbits are non-circular, the number of parameters that affect the trajectories now become dependent also on the respective eccentricities of the two starting orbits, the angle between their arguments of perigee, and for the non-coplanar problem the angle between the longitude of the ascending nodes. We shall examine the fuel-optimal maneuvers between two elliptic orbits with respect to variations only in the inclination of the two orbits for different time to rendezvous limits. For our example the interceptor initially resides on an elliptic orbit with classical orbit parameters given by $a_i = 1.4$ DU, $e_i = 0.2$, with $i_i = 0°$, longitude of the ascending node $\Omega_i = 0°$, argument of the perigee $\omega_i = 0°$. The target initial conditions are given by: $a_T = 2.0$DU, $e_T = 0.1$, longitude of the ascending node $\Omega_T = 0°$, argument of the perigee $\omega_i = 45°$. We shall examine the cost function variation corresponding to the target inclinations of $i_T = 0, 0.2, 0.5, 1.0$ radians. The variation of the characteristic velocity

with respect to the time to rendezvous are given in Figures 42-45. for the different target inclinations. Sample trajectories are shown for the target inclination of 0.2 radians and 1.0 radians in figures 46-51.

It may be worthwhile to explain how the trajectories were obtained. The two homotopy maps used were the same as the ones used previously for the circle-to-circle rendezvous problem. The solutions were obtained by systematically varying the orbit radii to obtain a circle-to-circle problem with the interceptor radius = 1.4DU, and target radius = 2.0DU for some time-of-flight limit. The circle-to-circle solution obtained is then used as the trivial solution in the same homotopy maps with the elliptic case providing the desired problem.

Ideally, one would use any circle-to-circle problem of section 5.4.1 and the ellipse-to-ellipse case as the trivial and the desired problem, respectively, in the homotopy map, map 2, given by equations [3.3.9]-[3.3.11] to yield the required solution. From a practical standpoint, however, this approach often fails to produce the desired result, since the resulting homotopy path does not meet the criteria required by the Chow-Yorke algorithm, i.e., it was not possible to show that the homotopy zero curve lies within some finite cylinder about the σ axis. Also, there may exist several solutions for the same starting value of the homotopy parameter. Yet another somewhat serious drawback is that the domain of our variables $Y = \{x, \lambda, \mu\}$, where $x = \{(\eta, \Delta u', \Delta h)_i, \eta_T\}$, $i = 1, \ldots,$ nim, and nim is the number of impulses and λ and μ are the Lagrange multipliers associated with the constraints, is not open in $\mathcal{R}^{n+ne+ni}$. Consequently, in tracking the homotopy curve since the inequality constraints are not enforced until the very end, often physically impossible regions are sought by the algorithm. It may for instance require a transfer angle greater than twice the asymptote

angle for a hyperbolic orbit, or require a negative transfer angle on a coasting arc. Thus the very advantage of using the homotopy map works against us if certain constraints have to be satisfied regardless of the path. Thus the fault lies mostly in not choosing a homotopy map which obviates the above mentioned physical impossibilities. Unfortunately, such a homotopy map has not yet been found, and to find one such provides ample scope for further research. The only alternative available then was to personally monitor the homotopy path so as to avoid the above mentioned pitfalls. On the other hand, using the equations [3.3.4]-[3.3.6], i.e., the Kuhn-Tucker equations themselves has the advantage that such pitfalls are easily avoided because of the inequality constraints, given by [4.3.12] as long as the active constraint set remains unchanged.

6.0 Concluding Remarks

Homotopy methods have been applied fairly successfully to parameter optimization problems, specifically to finding minimum-fuel solutions to the impulse thrust version of Lawden's problem. The principal idea is to track the solutions of a family of systems of equations formed by continuously deforming a trivial system to the desired system. At the culmination of the tracking procedure, the solution obtained satisfies the first order necessary conditions for the desired problem. Since a system of equations are required, the complementarity conditions on the inequality constraints are represented using Mangasarian's complementarity function. A second homotopy map was devised to eliminate singularities which occur at the transition point where some inequality constraint changes its active or inactive status.

The equations of motion of the satellite are expressed in terms of Burdett oscillator variables. Such a representation results in the simplification of the cost and constraint functions of the parameter optimization problem. It additionally facilitates obtaining the first and second partial derivatives required for tracking the homotopy zero curve. A homotopy map specific to the space-flight rendezvous problem is constructed, which

provides a valid space-flight problem for all values of the homotopy parameter between 0 and 1.

To obviate any trajectories 'cutting through the Earth,' a minimum radius constraint is considered, which appears as an infinite constraint. Polak's max-function approach is used to represent the minimum radius constraint. Since the max-function is not a C^2 function as required by the Chow-Yorke algorithm, a C^2 function is devised which can simulate the above mentioned max-function as closely as desired.

Several coplanar and non-coplanar solutions were obtained for both circle-to-circle and ellipse-to-ellipse cases. The coplanar solutions typically fall under two categories, one going inward to achieve the proper phase angle with the target for a Hohmann-like transfer. The other solution involves an outward thrust in an attempt to simulate the bi-elliptic transfer.

The non-coplanar case provides a wider variety of locally optimal solutions. Typically, if the inclinations are greater than about 0.5 radians, one would expect the Marchal-type transfer to be fuel-optimal, and depending on the time of flight restriction, the position where the principal out-of-plane impulse is applied would be different. However, for smaller target inclinations with respect to the interceptor orbit, global optimality of these transfers depends on the starting conditions of the target. For most target starting conditions, it is found to be more fuel-efficient to perform the principal out-of-plane maneuver at the site of intercept.

The present method thus provides an efficient procedure to obtain the variation of some system characteristic such as the cost function with respect to some system parameter

such as the time of flight using map 1, with map 2 used to bridge the singular situation. Map 2 may be used in itself to seek the solutions to some problem of interest.

The maps presented here are still not devoid of deficiencies. If the map 2 is used, the constraints are not satisfied until the very end of the curve tracking procedure, consequently, it is possible for the method to seek cost and constraint function values for variables for which these functions may not be defined. Also there is no guarantee that the solution curve is bounded within a finite cylinder about the σ axis. All this leaves ample room for further research to seek an even better map than the one provided, which satisfies all the requirements of the Chow-Yorke algorithm and yet is defined for all values of the variables that the algorithm may seek.

References

1. Allgower, E. and Georg, K., "Simplicial and continuation methods for approximating fixed points and solutions to systems of equations," *SIAM Review*, Vol 22, pp 28-85, 1980.

2. Bate, R.R., Mueller, D.D. and White, J.E., *Fundamentals of Astrodynamics.*,Dover Publications, Inc.,New York, 1971.

3. Bellman, R.E. and Wing, G.M., *An Introduction to Invariant Imbedding*, John Wiley & Sons, Inc., New York, 1975.

4. Bruce, J.W. and Giblin, P.J., *Curves and Singularities*, Cambridge Univeristy Press, New York, 1984.

5. Burdet, C.A., "Regularization of the two body problem, " *Zeitschrift für Angewandte Mathematik und Physik* , Vol 18, pp 434-438, 1967.

6. Chiu, J.-H., *Optimal Multiple-Impulse Nonlinear Orbital Rendezvous*, Ph.D. Thesis, University of Illinois at Urbana-Champaign, 1984.

7. Chiu, J.-H. and Prussing, J.E., "Optimal multiple-impulse time-fixed rendezvous between circular orbits," *Journal of Guidance, Control and Dynamics*, Vol 9, pp 17-22, 1986.

8. Chow, S.-N., Mallet-Paret, J. and Yorke, J.A., " Finding zeroes of maps: homotopy methods that are constructive with probability one," *Mathematics of Computation*, Vol 32, pp 887-899, 1978.

9. Clohessy, W.H. and Wiltshire, R.S., "Terminal guidance system for satellite rendezvous," *J Aerosp. Sc.*, Vol 27, pp 653-658, 1960.

10. Davidenko, D.F., "On the approximate solutions of systems of non-linear equation," *Ukrain. Mat. Zh.*, Vol 5, pp 196-206, 1953.

11. Deuflhard, P., "A stepsize control for continuation methods and its special application to multiple shooting techniques," *Numer. Math*, Vol 33, pp 115-146, 1979.

12. Eaves, B.C., "Homotopies for the computation of fixed points," *Math. Progamming*, Vol 3, pp 1-22, 1972.

13. Eaves, B.C. and Saigal, R., "Homotopies for computation of fixed points on unbounded regions," *Math. Progamming*, Vol 3, pp 225-237, 1972.

14. Edelbaum, T.N.,"How many impulses ?," *Astronautics and Aeronautics*, Vol 5, pp 64-69, Nov 1967.

15. Ficken, F.A., "The continuation method for functional equations," *Comm. Pure Appl. Math.*, Vol 4, pp 435-456, 1951.

16. Gfrerer, H., Guddat, J. and Wacker, Hj., "A globally convergent algorithm based on imbedding and parametric optimization," *computing*, Vol 30, pp 225-252, 1983.

17. Glandorf, D.R., *Optimal Fixed-Time Orbital Transfer with a Radial Constraint*, Ph.D. Thesis, Iowa State University, 1967.

18. Gobetz, F.W.,"Optimum transfers between hyperbolic asymptotes," *AIAA Journal*, Vol 1, pp 2034-2041, 1969.

19. Gobetz, F.W. and Doll, J.R.,"A survey of impulsive trajectories," *AIAA Journal*, Vol 5, pp 801-834, 1969.

20. Guillemin, V. and Pollack, A., *Differential Topology* , Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1974.

21. Hoelker, R.F. and Silber, R., "The bi-elliptical transfer between co-planar circular orbits," *Proc. 4th Symposium Ballistic Missile and Space Technology*, Los Angeles, 1959.

22. Keller, H.B., "Numerical solutions of bifurcation and nonlinear eigenvalue problems," in P. Rabinowitz, ed., *Applications of Bifurcation Theory*, Academic Press, New York, pp 359-384, 1977.

23. Kellogg, R.B., Li, T.Y. and Yorke, J., "A constructive proof of the Brouwer fixed-point theorem and computational results," *SIAM J Numer Anal.*, Vol 13, pp 473-483, 1976.

24. Klopfenstein, R.W., "Zeros of nonlinear functions," *J Assoc. Comput. Mech.* , Vol 8, pp 336-373, 1961.

25. Kubicek, M., "Dependence of solutions of nonlinear equations on a parameter," *ACM Trans. Math. Software* , Vol 2, pp 98-107, 1976.

26. Lawden, D.F., "Interplanetary rocket trajectories.," *Advances in Space Sciences*, Ed. by F.I. Ordway, Chap. 1, Academic Press, New York, 1959.

27. Lawden, D.F., "Optimal intermediate thrust arcs in a gravitational field.," *Astronautica Acta*, Vol 8, pp 106-123, 1962.

28. Lawden, D.F., *Optimal Trajectories for Space Navigation*, Butterworth & Co. (Publishers) Ltd., London, 1963.

29. Lefton, L. and Kelley, H.J., *Variable-metric projection optimization program notes*, Report No. 77-12, Contract NAS 1-14891, July 1978.

30. Lion, P.M. and Handelsmann, M., "Primer vector on a fixed-time impulsive trajectories," *AIAA Journal*, Vol 6, pp 127-132, 1968.

31. Lutze, F.H., Cliff, E.M. and Kelley, H.J., *Optimal Rendezvous with Cooperative Vehicles II*, VPI&SU Report, July 1985.

32. Mangasarian, O.L. and Fromovitz, S., "The Fritz-John necessary optimality conditions in the presence of equality and inequality constraints," *J Math. Anal. and Appl.*, Vol 17, pp 37-47, 1967.

33. Mangasarian, O.L., "Equivalence of the complementarity problem to a system of non-linear equations," *SIAM J Appl. Math.*, Vol 31, pp 89-92, 1976.

34. Marec, J.P., *Trajectoires Spatiales Optimales*, Cours de l'ENSAE, Toulouse, 1973.

35. Marec, J.P., *Optimal Space Trajectories*, Elsevier, New York, New York, 1979.

36. Neustadt, L.W., "Optimization, A moment problem and non-linear programming," *SIAM Journal on Control*, Vol 2, pp 33-53, 1964.

37. Neustadt, L.W., "A general theory of minimum-fuel space trajectories," *SIAM Journal on Control*, Vol 3, pp 317-356, 1966.

38. Ortega, J.M. and Rheinboldt, W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, Inc., Orlando, Florida 32887, 1970.

39. Polak, E., "Notes on the Mathematical Foundations of Non-differentiable Optimization in Engineering Design," *Electronics Research Laboratory Memorandum No. UCB/ERL M84/15*, University of California, Berkeley, CA 94720, 1984.

40. Poore, A.B. and Al-Hassan, Q., "The expanded Lagrangian system for constrained optimization problems," *SIAM J Control and Opt.* , Vol 26, pp 417-427, 1988.

41. Potter, J.E. and Stern, R.E., "Optimization of midcourse velocity corrections," *Proceedings IFAC Symposium on Automatic Control in the Peaceful Uses of Space*, Stavanger, Norway, Plenum Press, N.Y., pp 70-84, June 1965.

42. Prussing, J.E., "Optimal four impulse fixed-time rendezvous in the vicinity of a circular orbit," *AIAA Journal*, Vol 7, pp 928-935, 1969.

43. Prussing, J.E., "Optimal two and three impulse fixed-time rendezvous in the vicinity of a circular orbit," *AIAA Journal*, Vol 8, pp 1221-1228, July 1970.

44. Robbins, H.M., "An analytical study of the impulsive approximation," *AIAA Journal*, Vol 4, pp 1417-1423, 1966.

45. Roberts, S.M. and Shipman, J.S., "Continuation in shooting methods for two-point boundary value problems," *J Math. Anal. Appl.*, Vol 21, pp 23-30, 1968.

46. Saigal, R. and Todd, M.J., "Efficient acceleration techniques for fixed point algoithms," *SIAM J Numer. Anal.*, Vol 15, pp 997-1007, 1978.

47. Scarf, H., "The approximation of fixed points of continuous mappings," *SIAM J Appl. Math.*, Vol 15, pp 1328-1343, 1967.

48. Schumacher, P.W., *Results of True-Anomaly Regularization in Orbital Mechanics*, Ph.D. Dissertation, Dept. of Aerospace and Ocean Engg., VPI & SU, 1987.

49. Scott, M.R., *Invariant Imbedding and its Applications to Ordinary Differential Equations, An Introduction.*, Addison- Wesley Publishing Co., Reading, Massachusetts, 1973.

50. Wacker, H., "A summary of the developments on imbedding methods," *Continuation Methods*, edited by H. Wacker, Academic Press, Inc., New York, pp 1-35, 1978.

51. Watson, L.T., "Solving the non-linear complementarity problem by a homotopy Method," *SIAM J Control and Optimization*, Vol 17, pp 36-46, 1979.

52. Watson, L.T., "A globally convergent algorithm for computing fixed points of C^2 maps," *Appl. Math. and Comput.*, Vol 5, pp 297-311, 1979.

53. Watson, L.T., "Engineering applications of the Chow-Yorke algorithm," *Appl. Math. and Comput.*, Vol. 9, pp 111-133, 1981.

54. Watson, L.T., "Numerical linear algebra aspects of globally convergent homotopy methods," *SIAM Review*, Vol. 28, pp 529-545, 1986.

55. Watson, L.T., Billups, S.C. and Morgan, A.P., "HOMPACK: A Suite of Codes for Globally Convergent Homotopy Algorithms," *GMRL Research Publication GMR-5344*, General Motors Research Labs., Warren, MI 48109, June 1986.

56. Watson, L.T., Bixler, J.P. and Poore, A.B., "Continuous homotopies for the linear complementarity problem," *Virginia Tech Computer Science Report, TR 87-38*, VPI & SU, Blacksburg, VA 24060, Dec. 1987.

57. Vasudevan, G., *Fuel Optimal Rendezvous including a Radial Constraint*, M.S. Thesis, Dept. of Aerospace and Ocean Engg., VPI & SU, 1986.

58. Vasudevan, G. and Lutze, F.H., "On verification of fuel-optimal space trajectories," *Opt. Cont. Appl. Meth.*, Vol 9., pp 285-302, 1988.

59. Yakovlev, M.N., "Solution of systems of nonlinear equations by the method of differentiation with respect to a parameter," *NASA TT F-254* , also "K resheniyu sistem nelineynykh uravneniy metodom differentsirovaniya po parametru, " *Zhurnal Vychislitel'noy Matematiki i Matematicheskoy Fiziki* , Vol 4, pp 146-149, 1964.

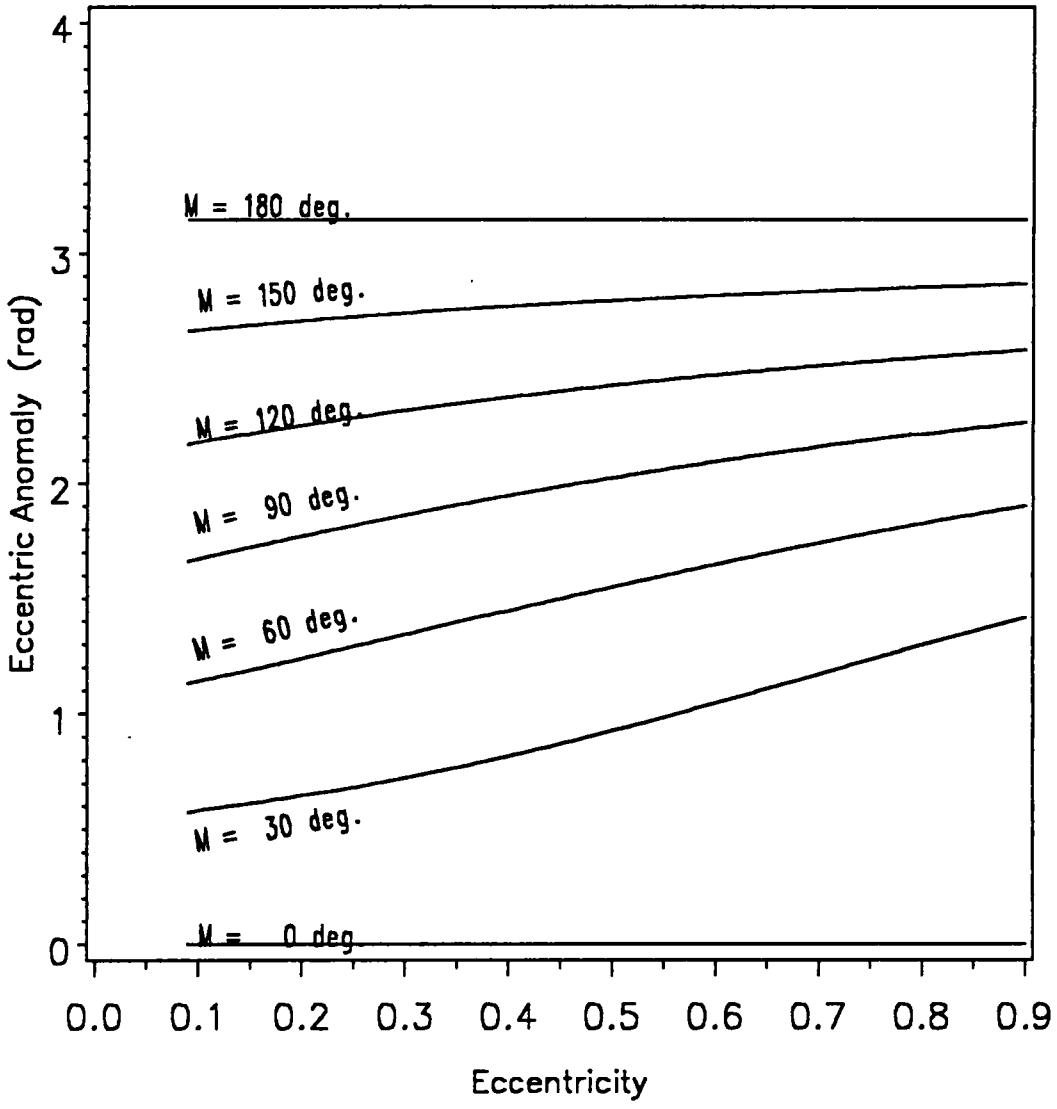

Figure 1. Eccentric anomaly vs. eccentricity.: Variation of the eccentric anomaly with respect to eccentricity corresponding to different values of the mean anomaly.

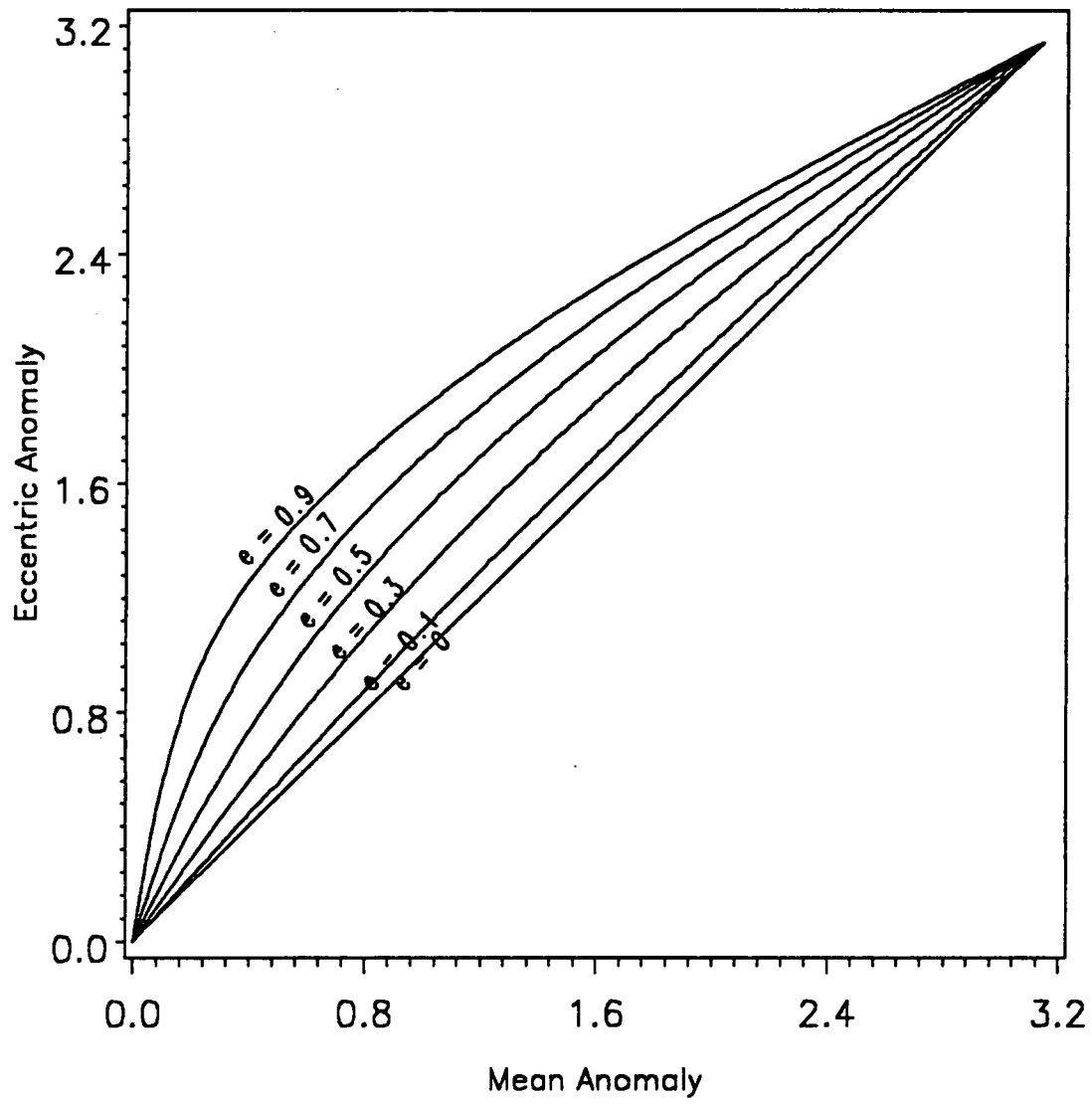

Figure 2. Eccentric anomaly vs. mean anomaly.: Variation of the eccentric anomaly with respect to mean anomaly corresponding to different values of the eccentricity.

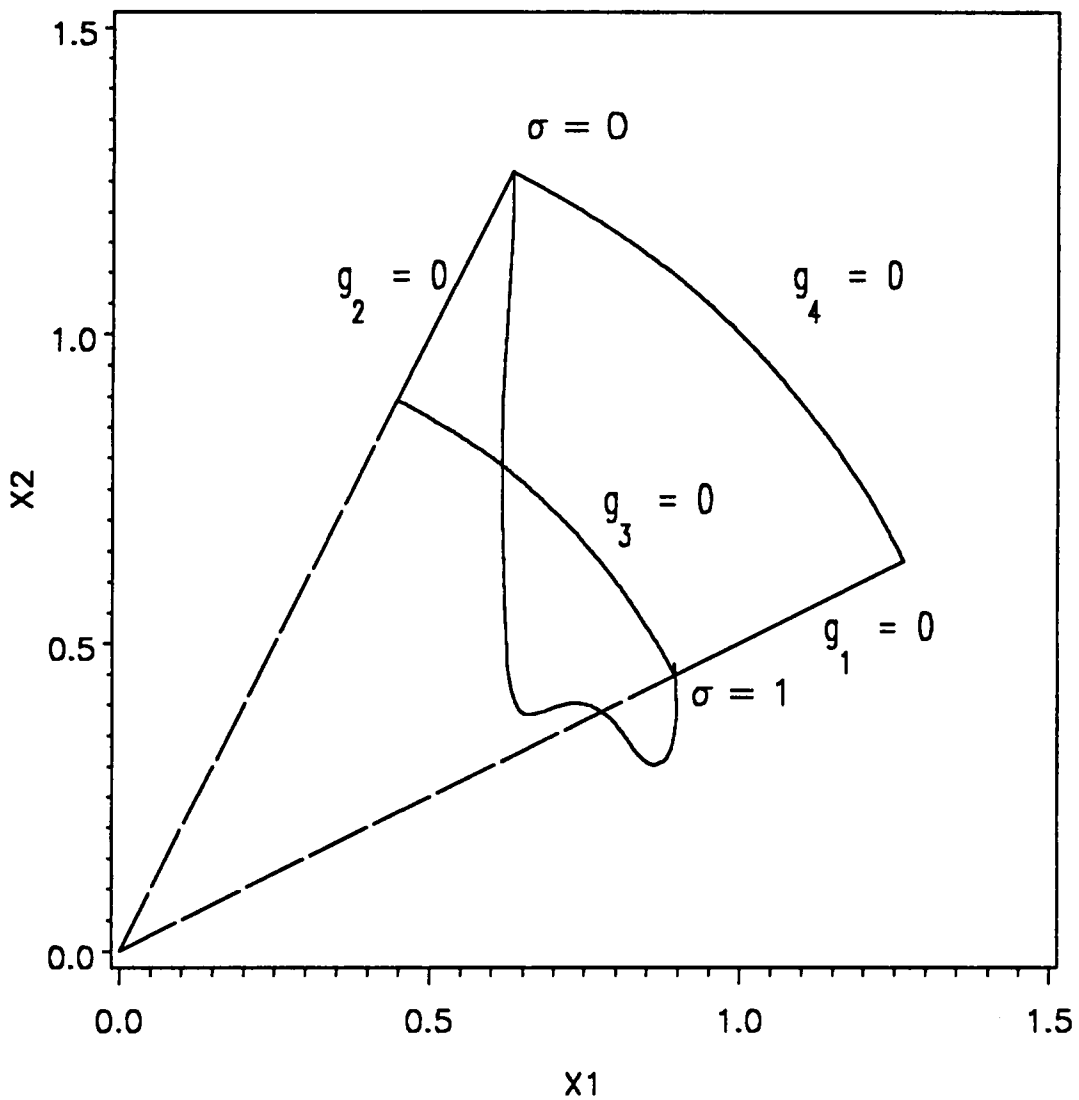

Figure 3. Feasible region for Example 1 and the plot of x_1 versus x_2 with respect to the variation in σ.

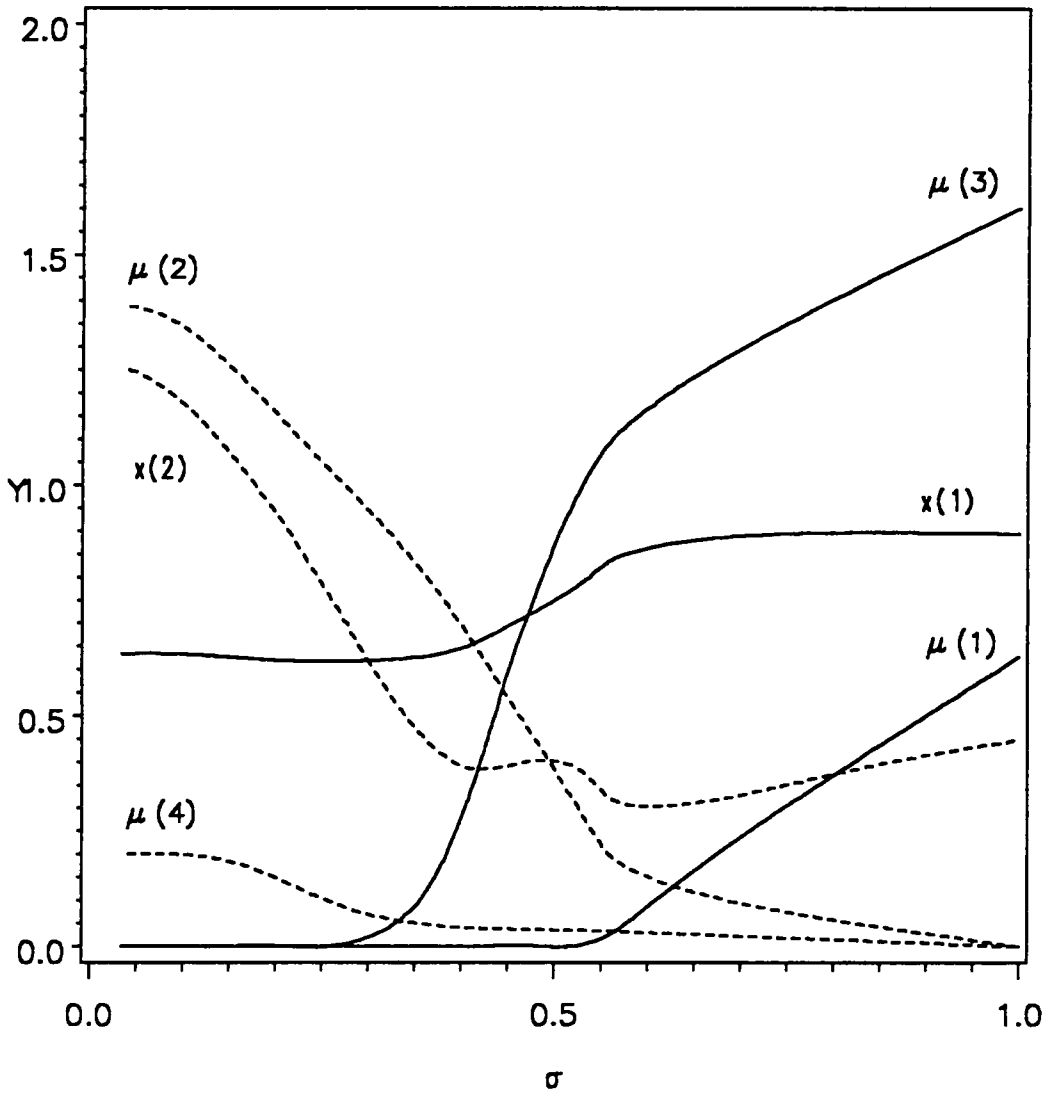

Figure 4. Variation of Y = $\{x_1, x_2, \mu_1, \mu_2, \mu_3, \mu_4\}$ with respect to σ.

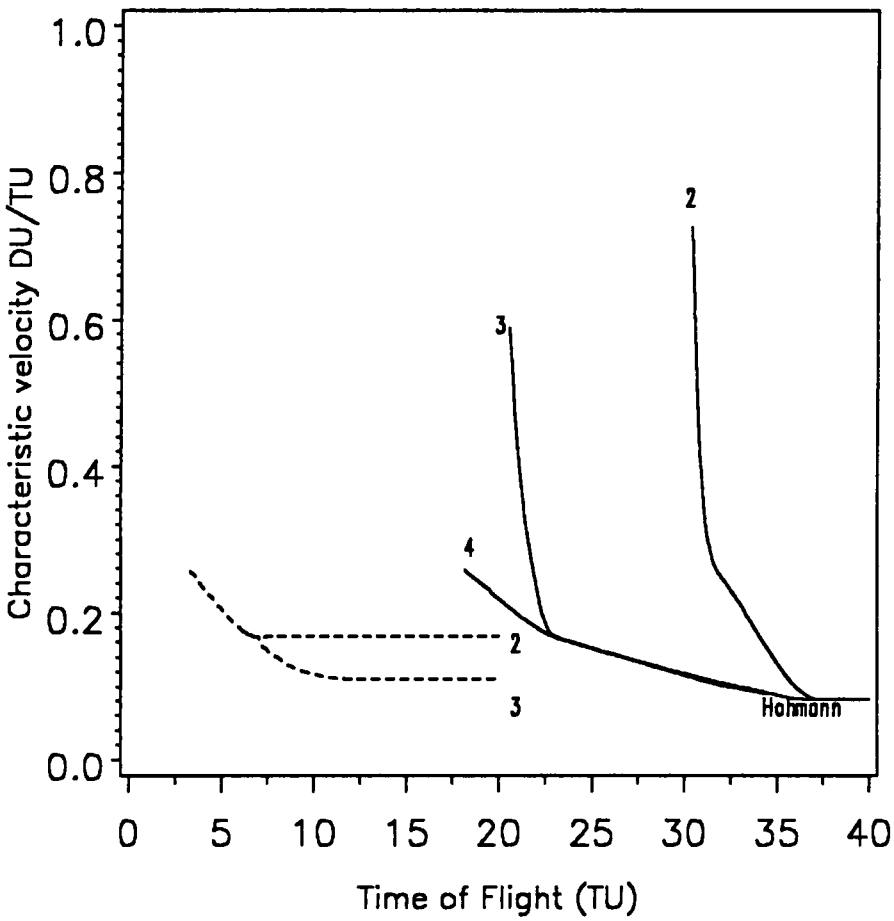

Figure 5. Characteristic velocity versus time to rendezvous for $\beta_T = 0°$.: the minimum radius limit = 1 DU., $i_T = 0°$

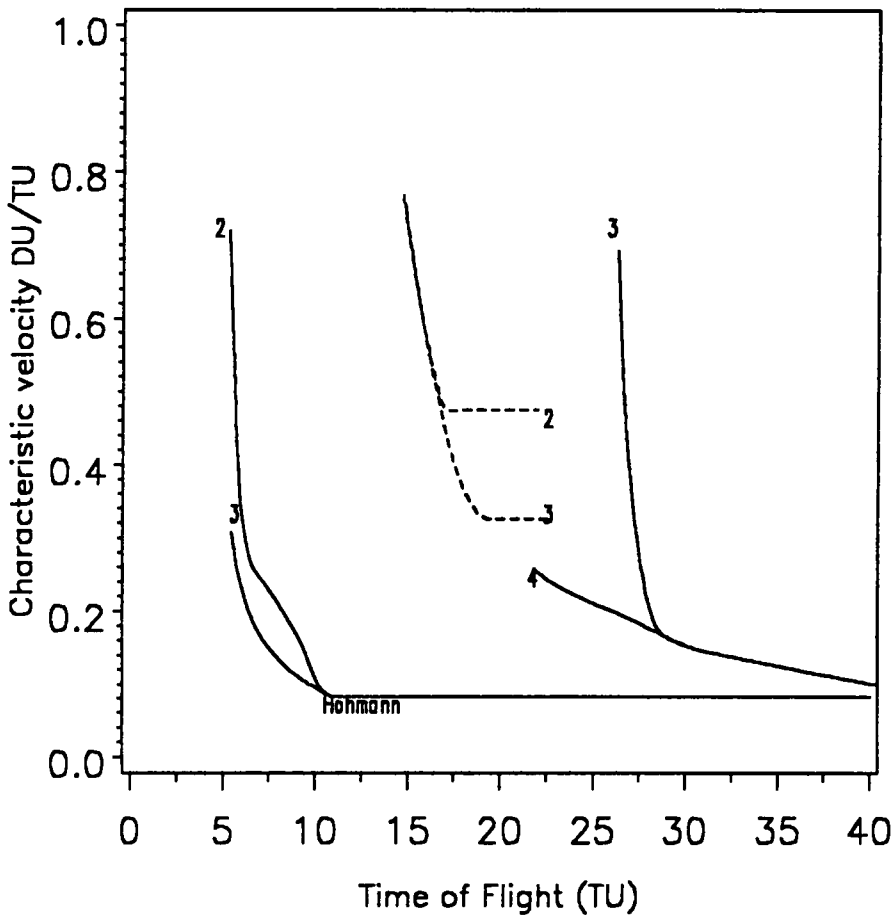

Figure 6. Characteristic velocity versus time to rendezvous for $\beta_T = 90°$.: the minimum radius limit = 1 DU., $i_T = 0°$

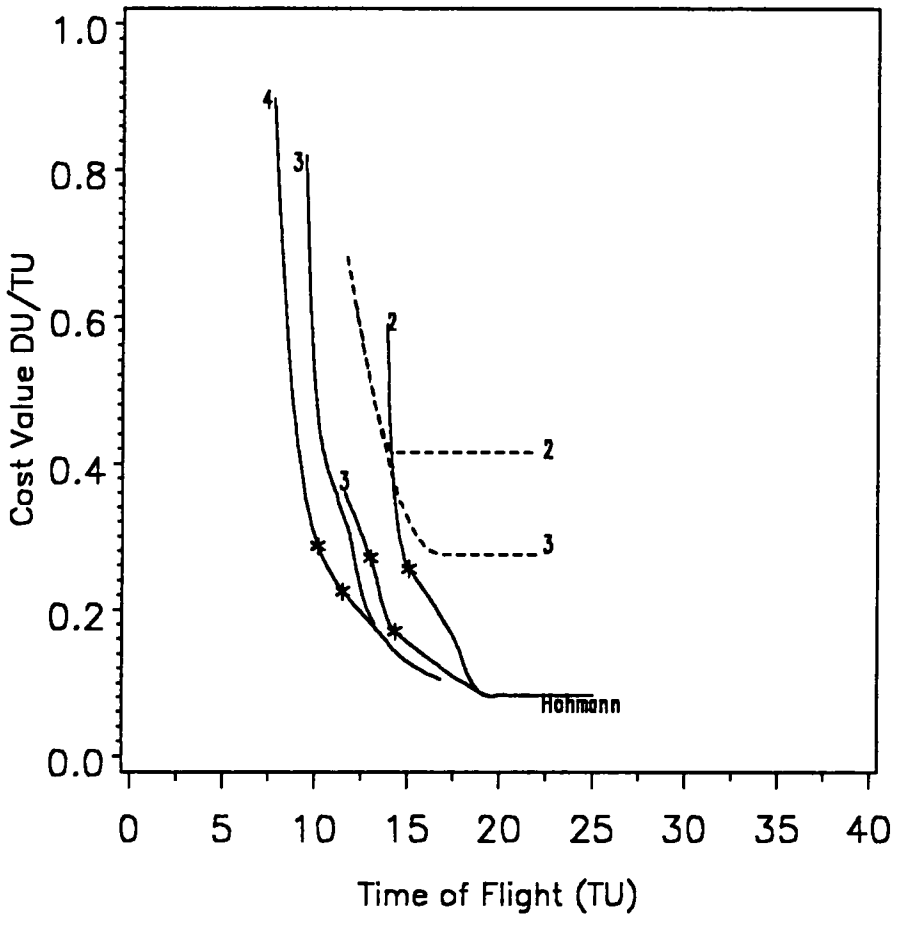

Figure 7. Characteristic velocity versus time to rendezvous for $\beta_T = 180°$.: the minimum radius limit = 1 DU., $i_T = 0°$

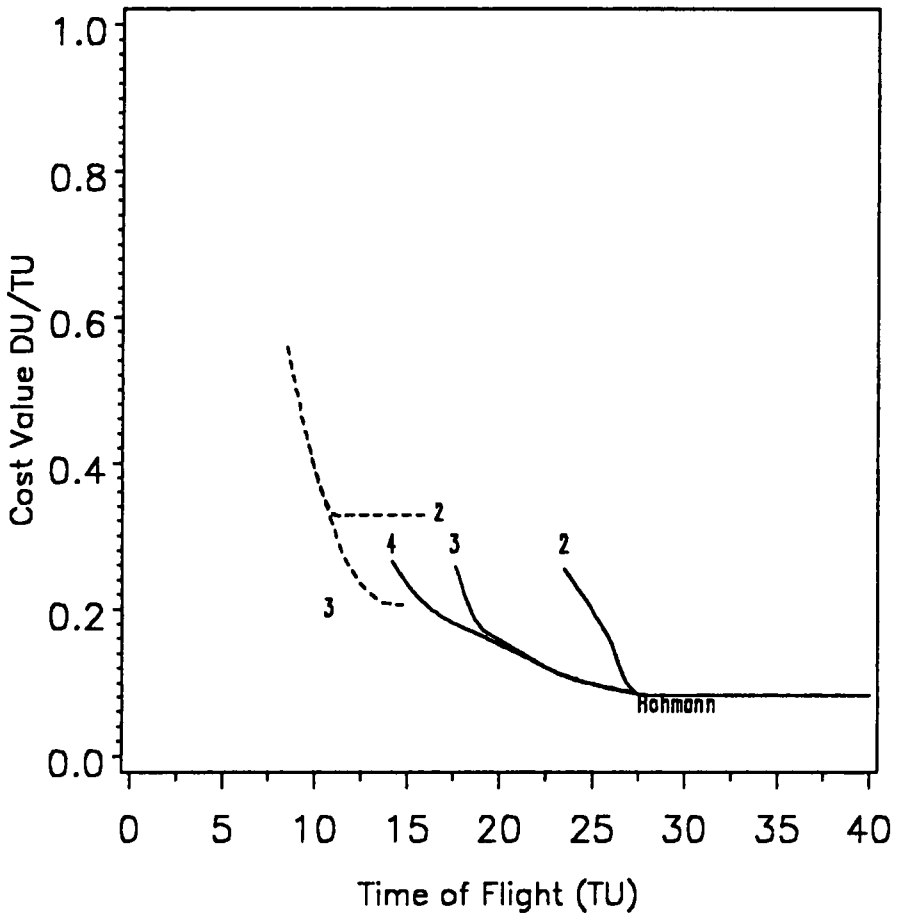

Figure 8. Characteristic velocity versus time to rendezvous for $\beta_T = 270°$.: the minimum radius limit = 1 DU., $i_T = 0°$

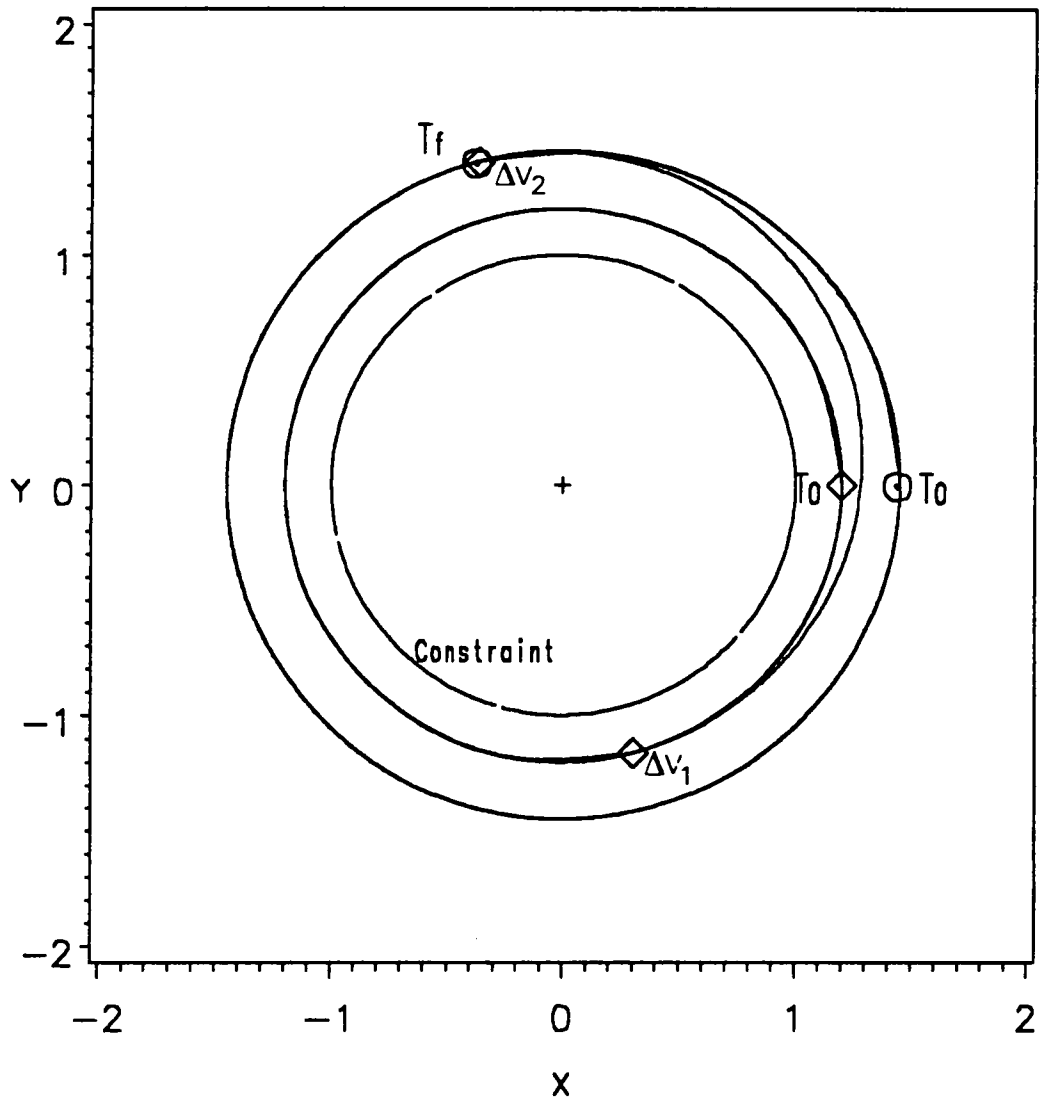

Figure 9. 2 impulse trajectory (Hohmann transfer).: Target phase at start $\beta_T = 0°$, target orbit inclination $i_T = 0°$, cost = 0.08226 DU/TU. Time of flight = 36.10 TU (485.42 minutes).

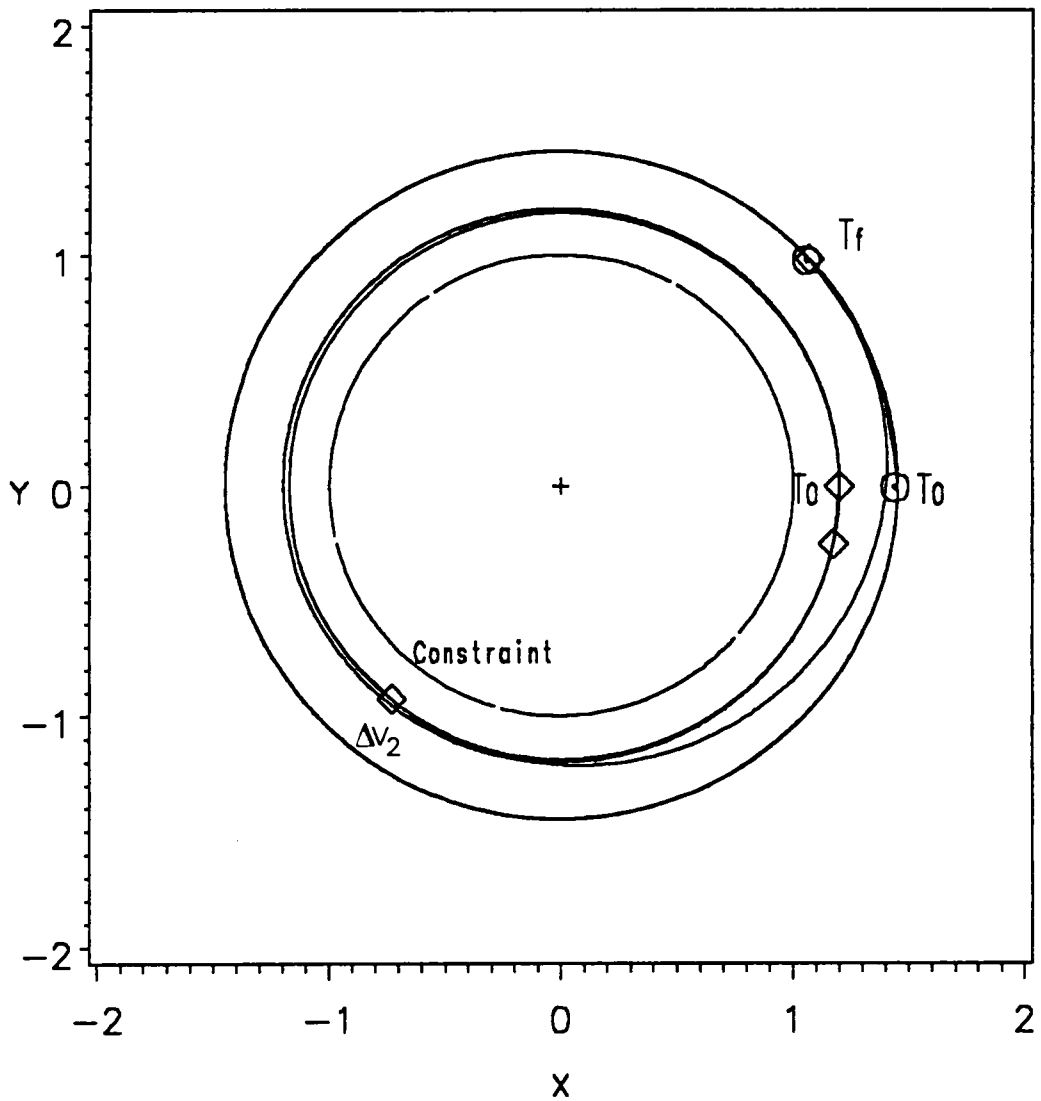

Figure 10. 3 impulse trajectory (type A).: Target phase at start $\beta_T = 0°$, target orbit inclination $i_T = 0°$, cost = 0.09288 DU/TU. Time of flight = 34.21 TU (460.0 minutes).

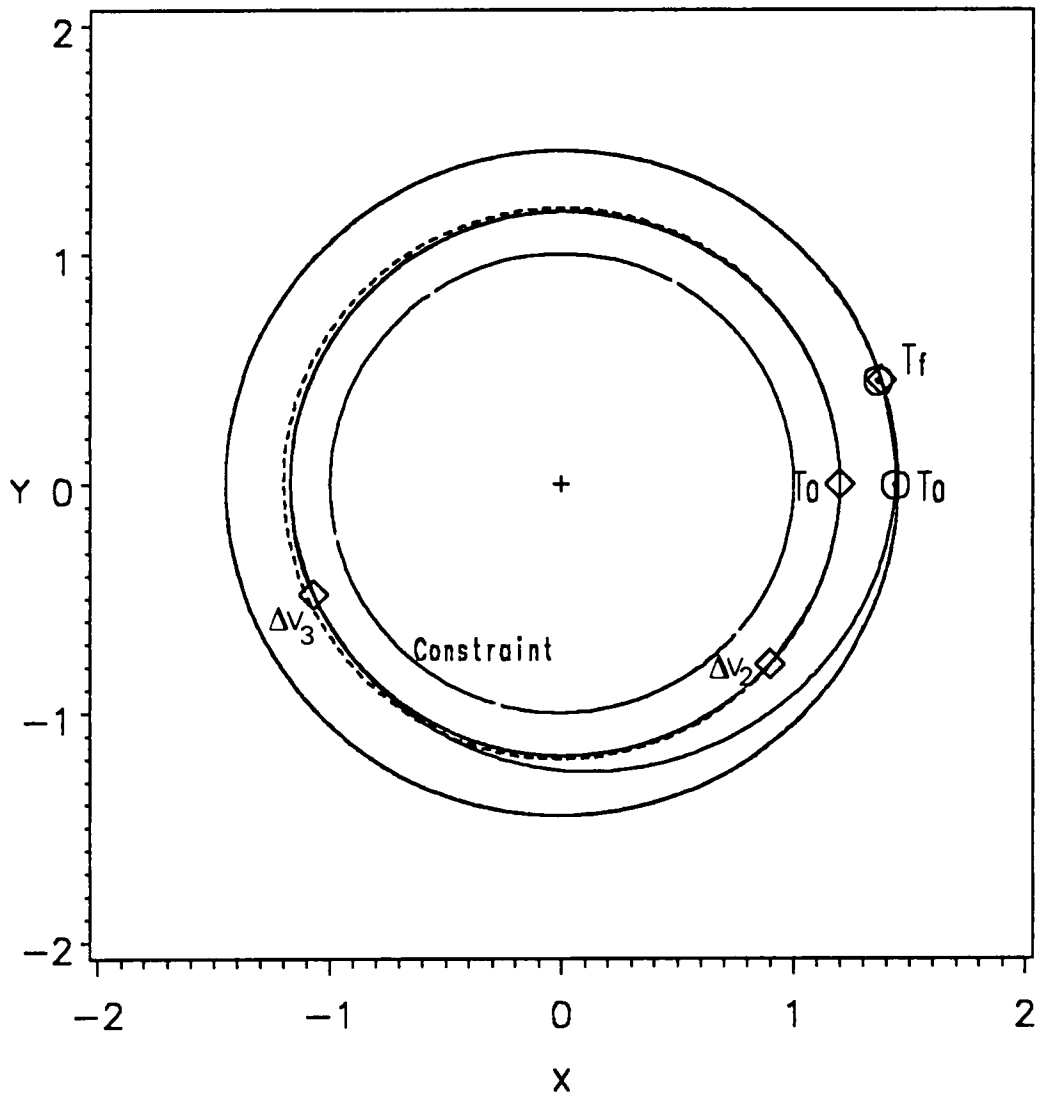

Figure 11. 4 impulse trajectory (type A).: Target phase at start $\beta_T = 0°$, target orbit inclination $i_T = 0°$, cost = 0.09472 DU/TU. Time of flight = 33.47 TU (450.0 minutes).

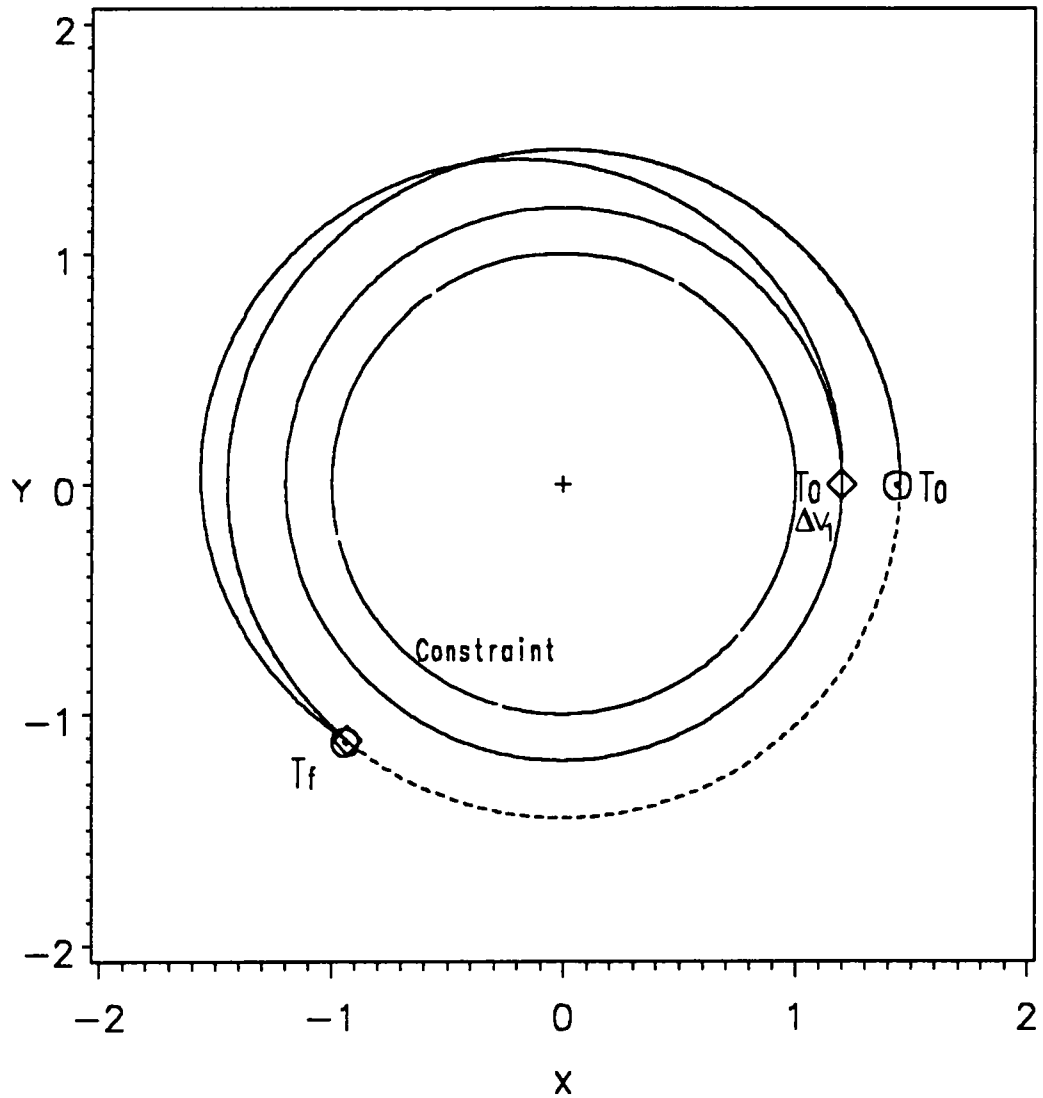

Figure 12. 2 impulse trajectory (type B).: Target phase at start $\beta_T = 0°$, target orbit inclination $i_T = 0°$, cost = 0.1676 DU/TU. Time of flight = 7.01 TU (94.30 minutes).

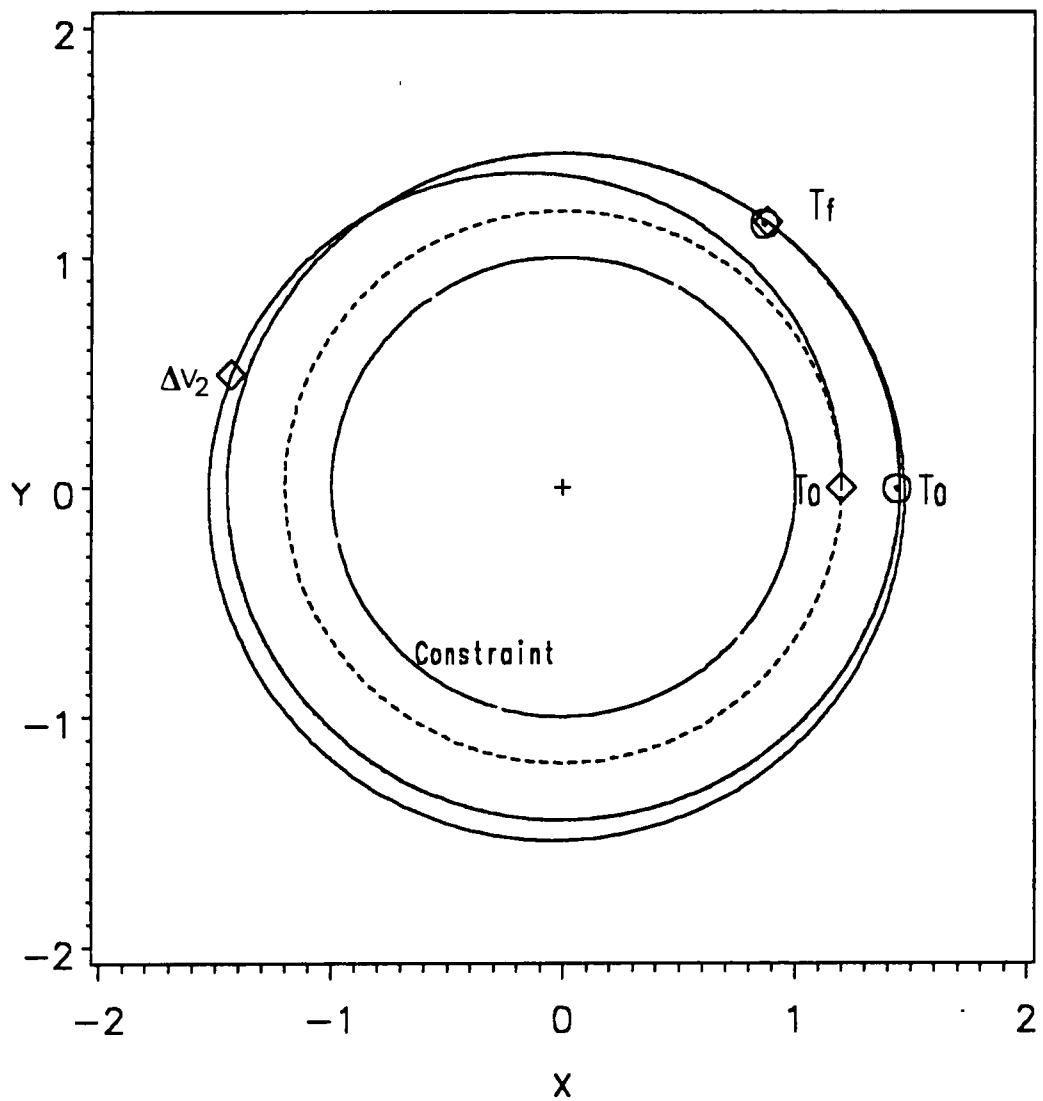

Figure 13. 3 impulse trajectory (type B).: Target phase at start $\beta_T = 0°$, target orbit inclination $i_T = 0°$, cost = 0.1103 DU/TU. Time of flight = 12.57 TU (169.09026 mins).

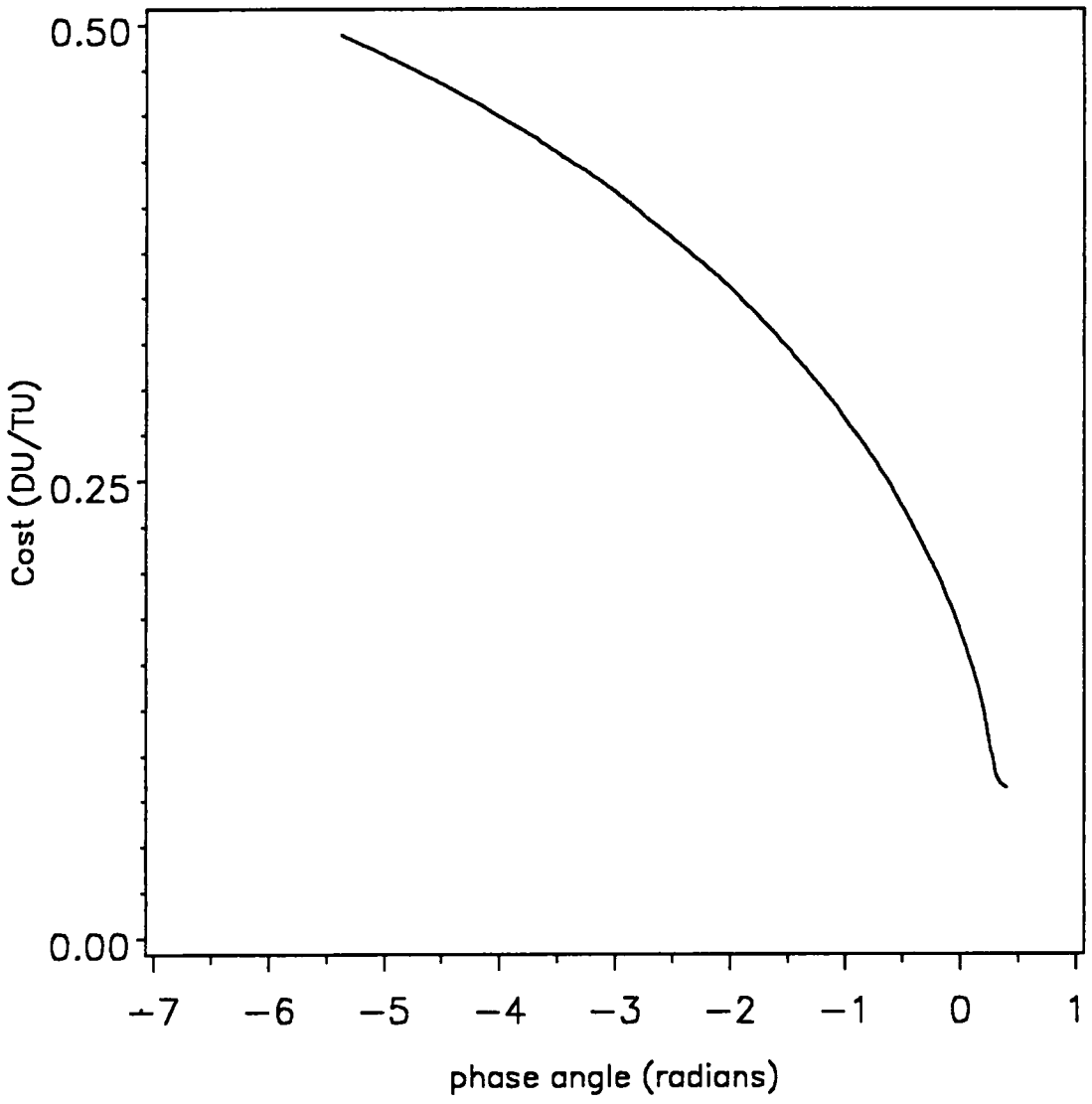

Figure 14. Characteristic velocity versus target phase angle (2 impulse): Number of impulses = 2. The time free rendezvous case, corresponding to the B type orbits.

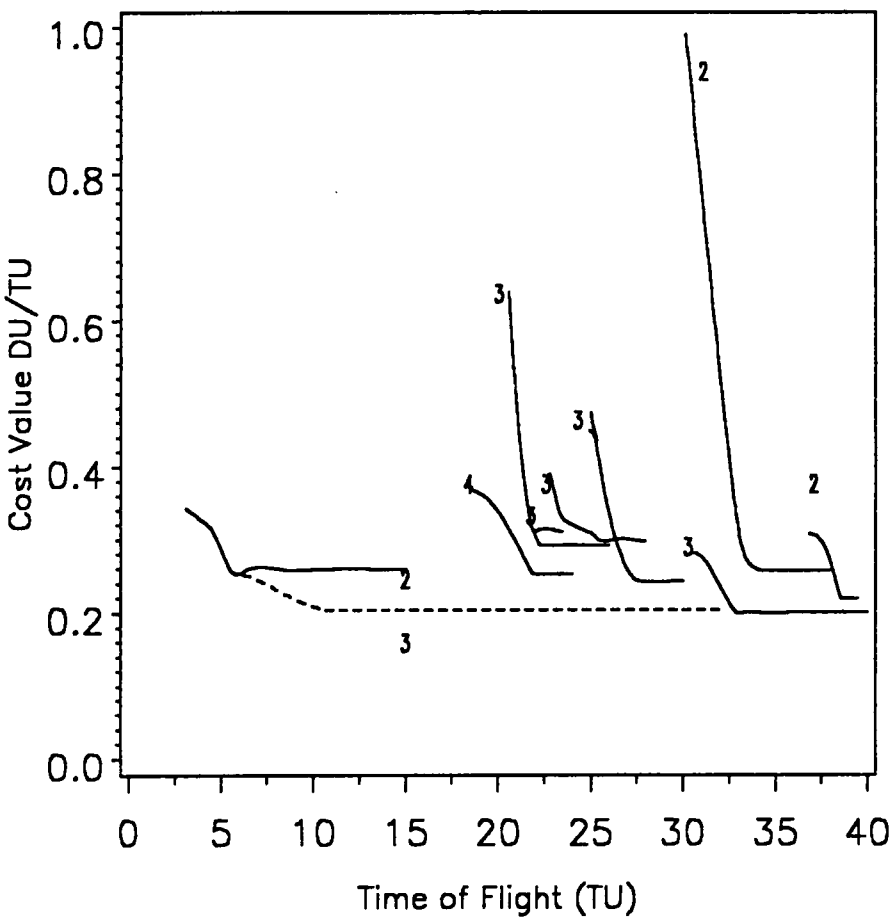

Figure 15. Characteristic velocity versus time to rendezvous for $\beta_T = 0°$, $i_T = 0.2$ radians.: the minimum radius limit = 1 DU.,

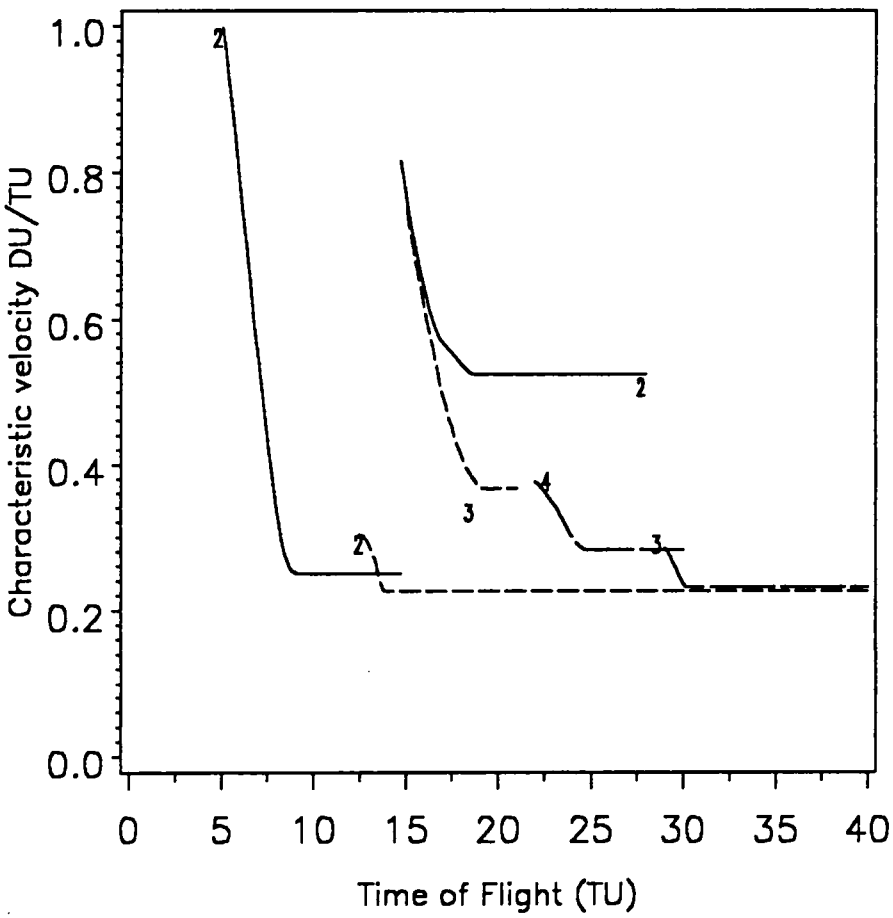

Figure 16. Characteristic velocity versus time to rendezvous for $\beta_T = 90°$, $i_T = 0.2$ radians.: the minimum radius limit = 1 DU.,

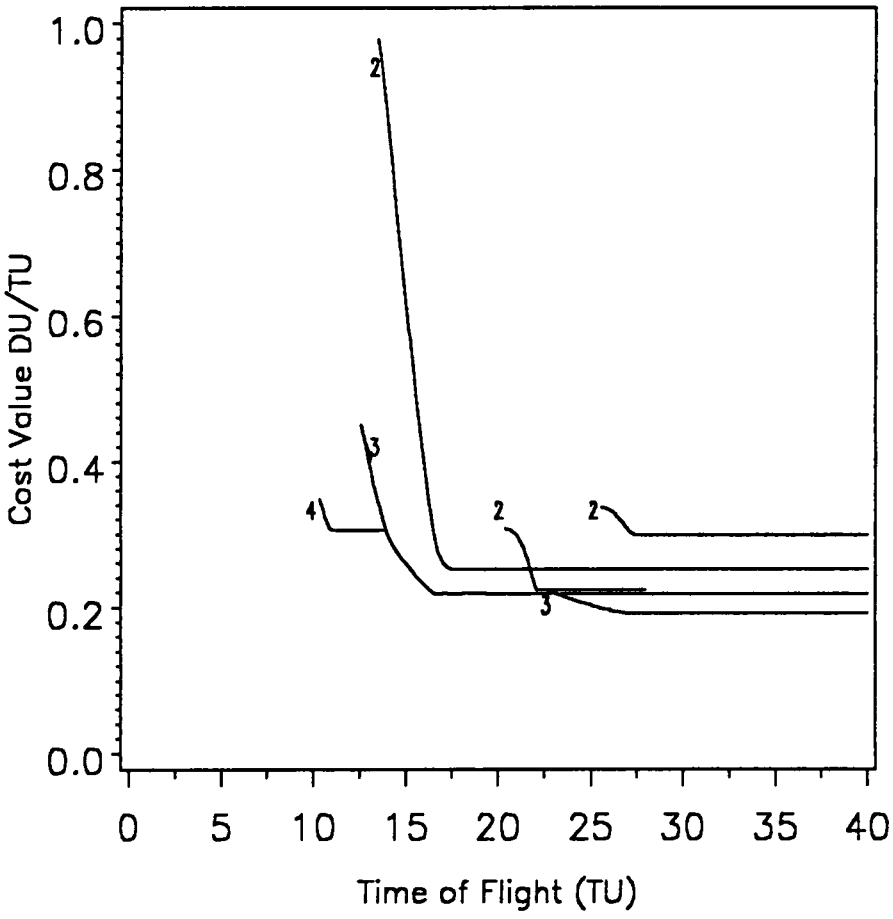

Figure 17. Characteristic velocity versus time to rendezvous for $\beta_T = 180°$, $i_T = 0.2$ radians: the minimum radius limit = 1 DU.,

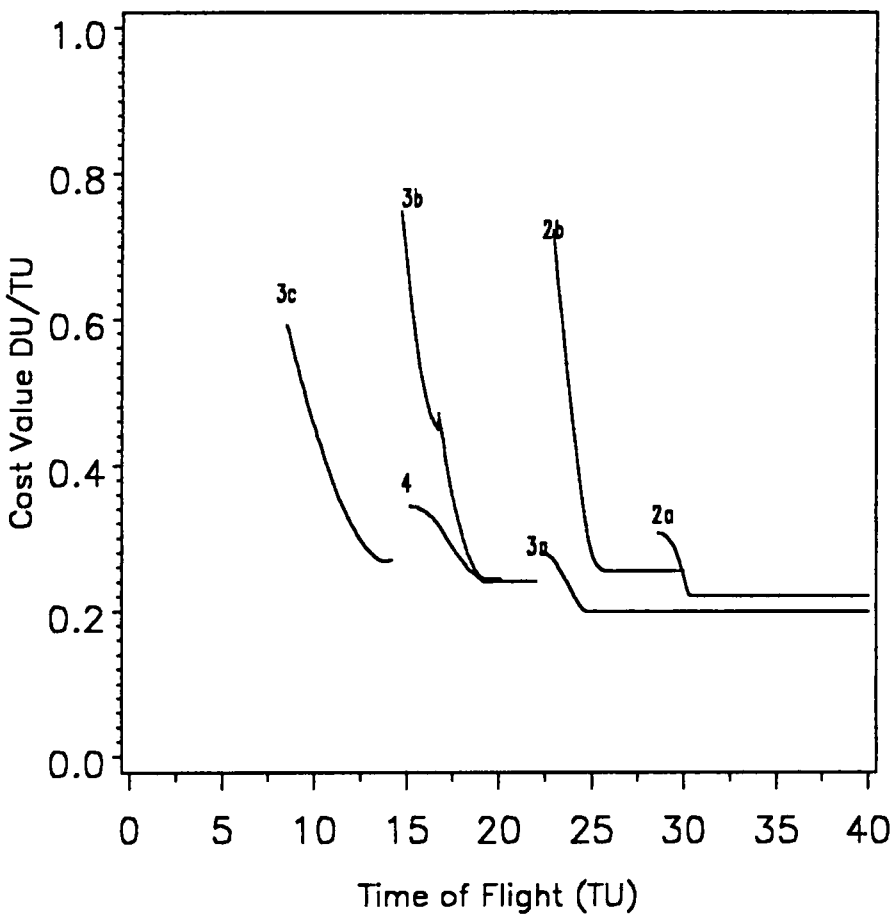

Figure 18. Characteristic velocity versus time to rendezvous for $\beta_T = 270°$, $i_T = 0.2$: the minimum radius limit = 1 DU., radians.

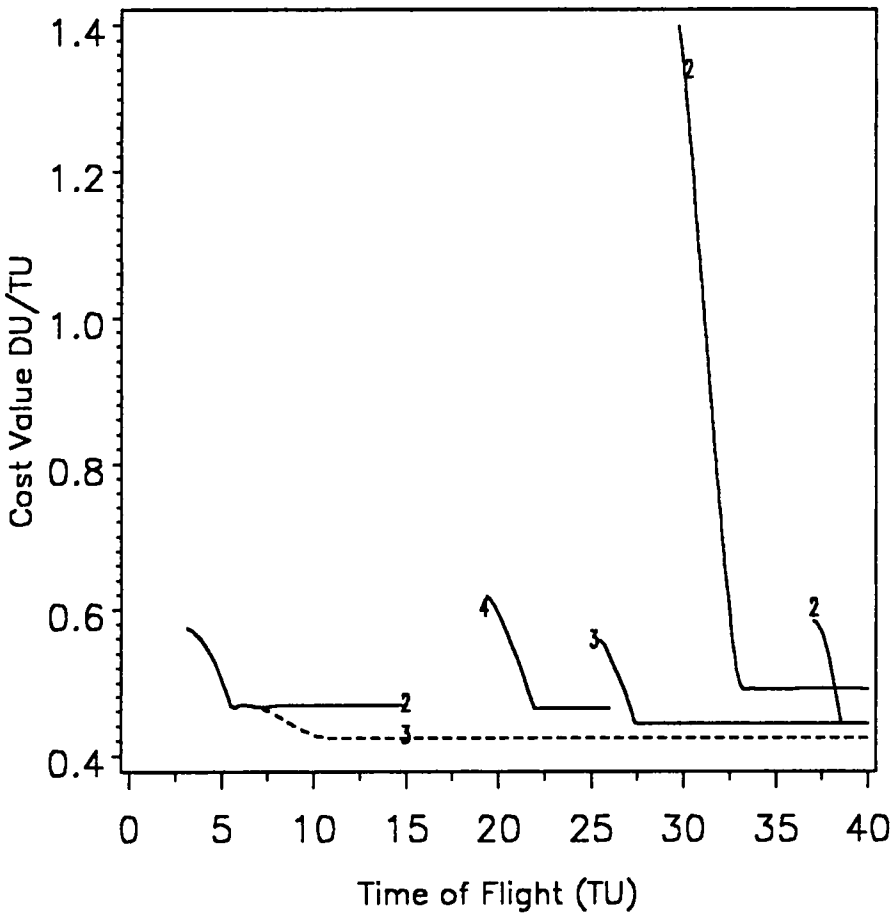

Figure 19. Characteristic velocity versus time to rendezvous for $\beta_T = 0°$, $i_T = 0.5$ radians.: the minimum radius limit = 1 DU.,

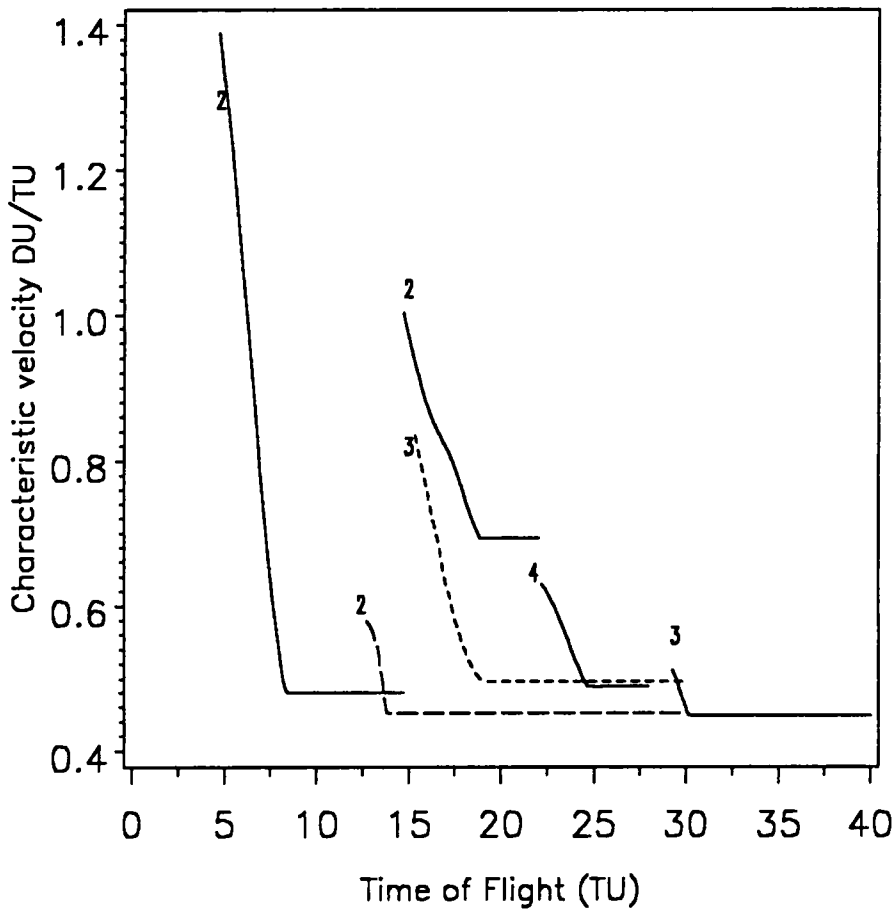

Figure 20. Characteristic velocity versus time to rendezvous for $\beta_T = 90°$, $i_T = 0.5$ radians.: the minimum radius limit = 1 DU.,

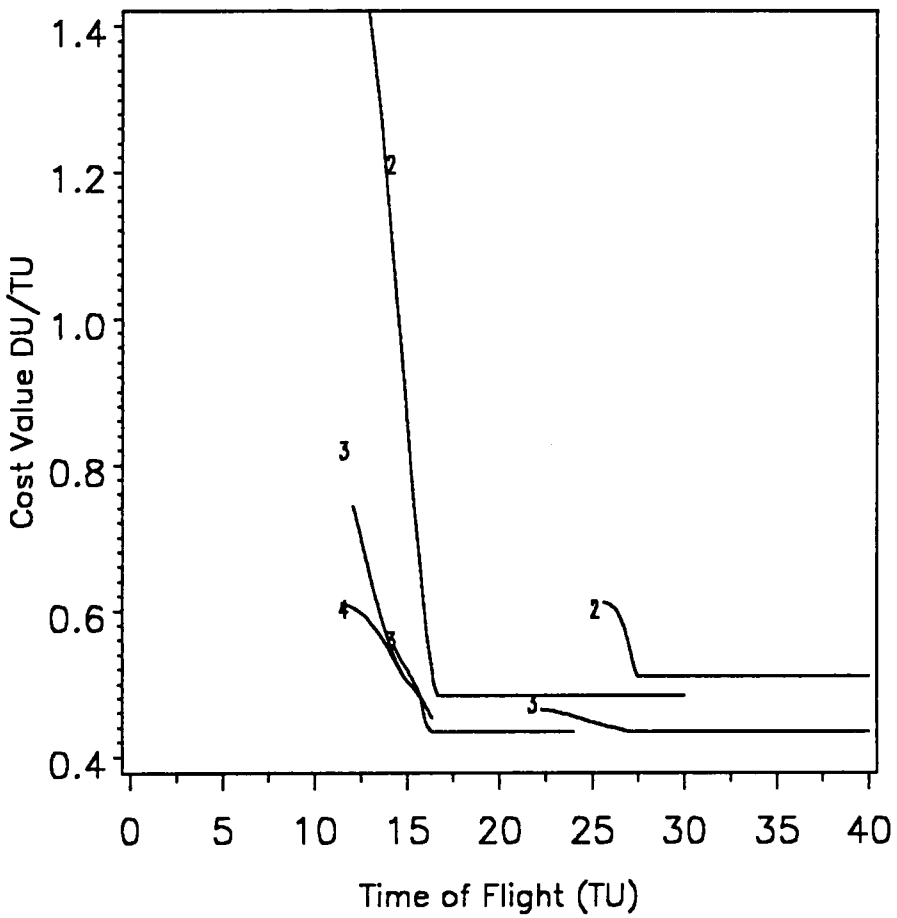

Figure 21. Characteristic velocity versus time to rendezvous for $\beta_T = 180°$, $i_T = 0.5$ radians.: the minimum radius limit = 1 DU.,

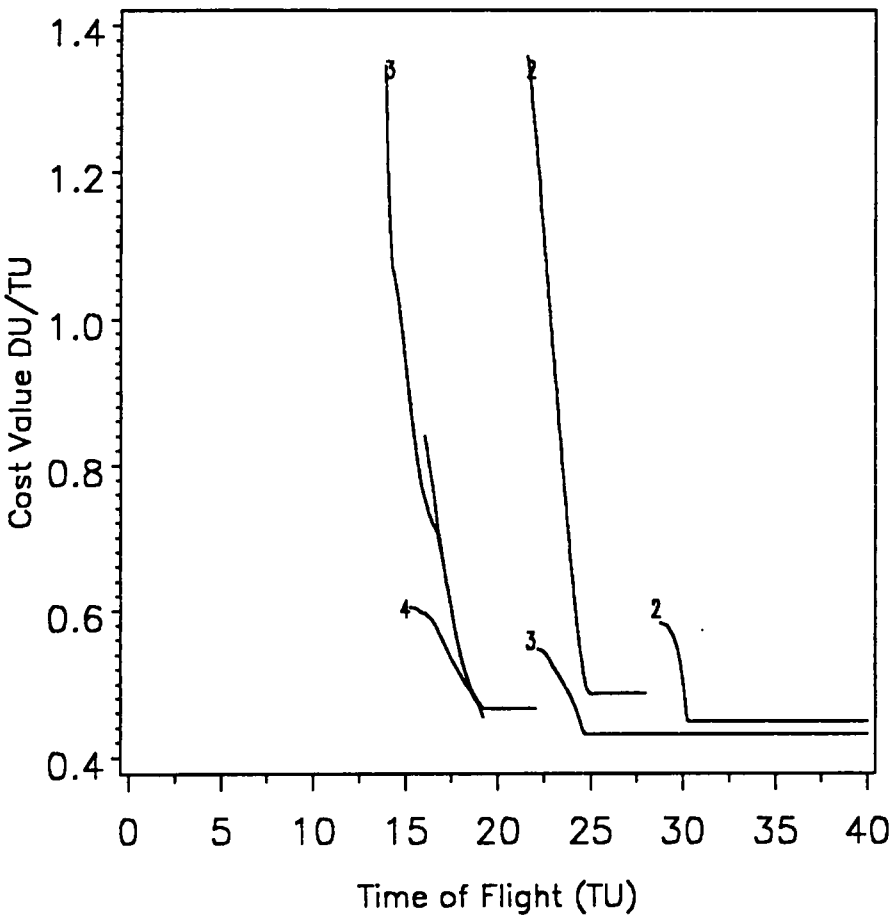

Figure 22. Characteristic velocity versus time to rendezvous for $\beta_T = 270°$, $i_T = 0.5$ radians.: the minimum radius limit = 1 DU.,

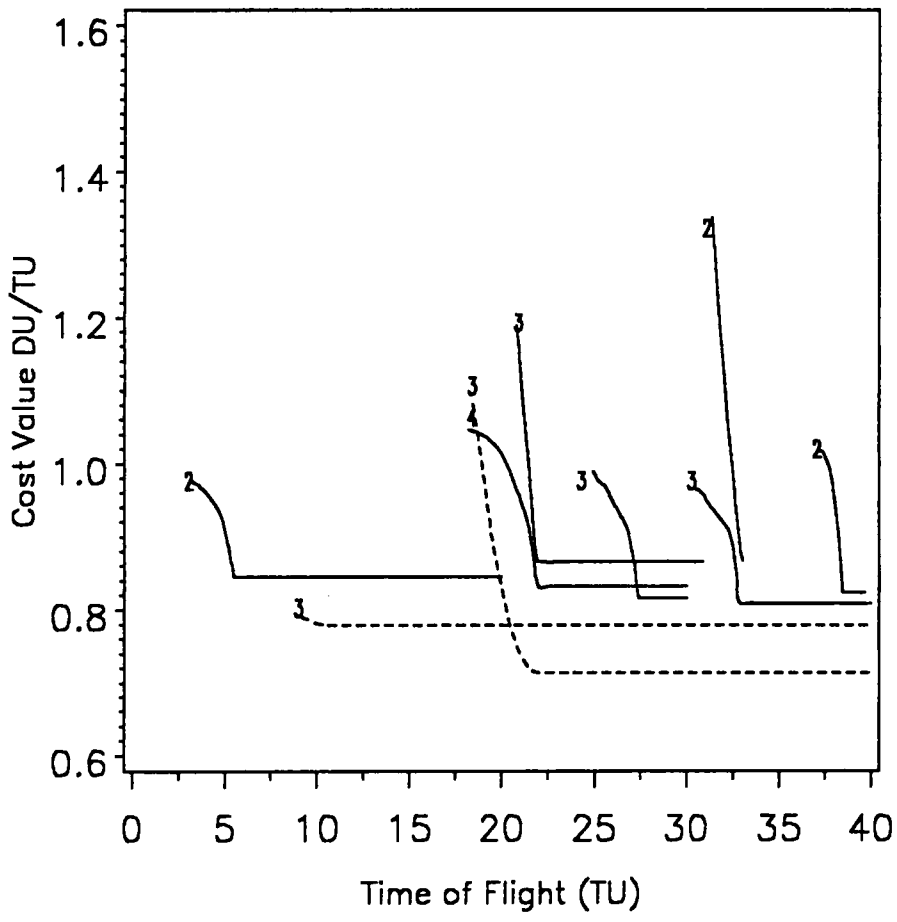

Figure 23. Characteristic velocity versus time to rendezvous for $\beta_T = 0°$, $i_T = 1$ radian.: the minimum radius limit = 1 DU.,

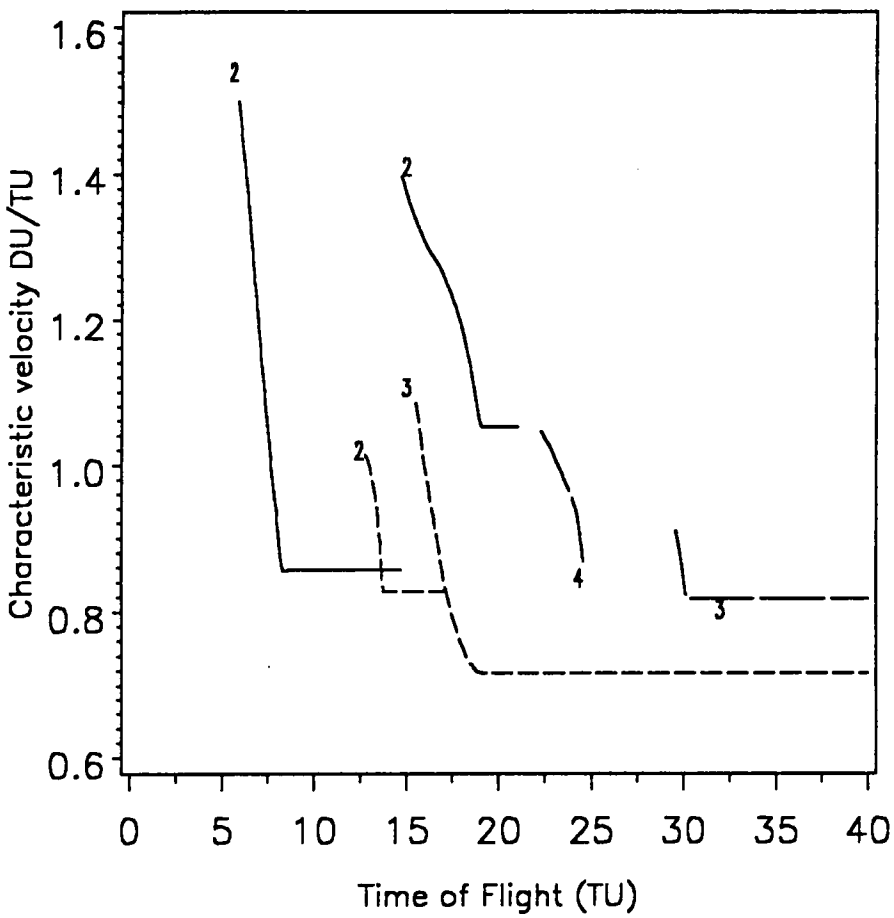

Figure 24. Characteristic velocity versus time to rendezvous for $\beta_T = 90°$, $i_T = 1$ radian.: the minimum radius limit = 1 DU.,

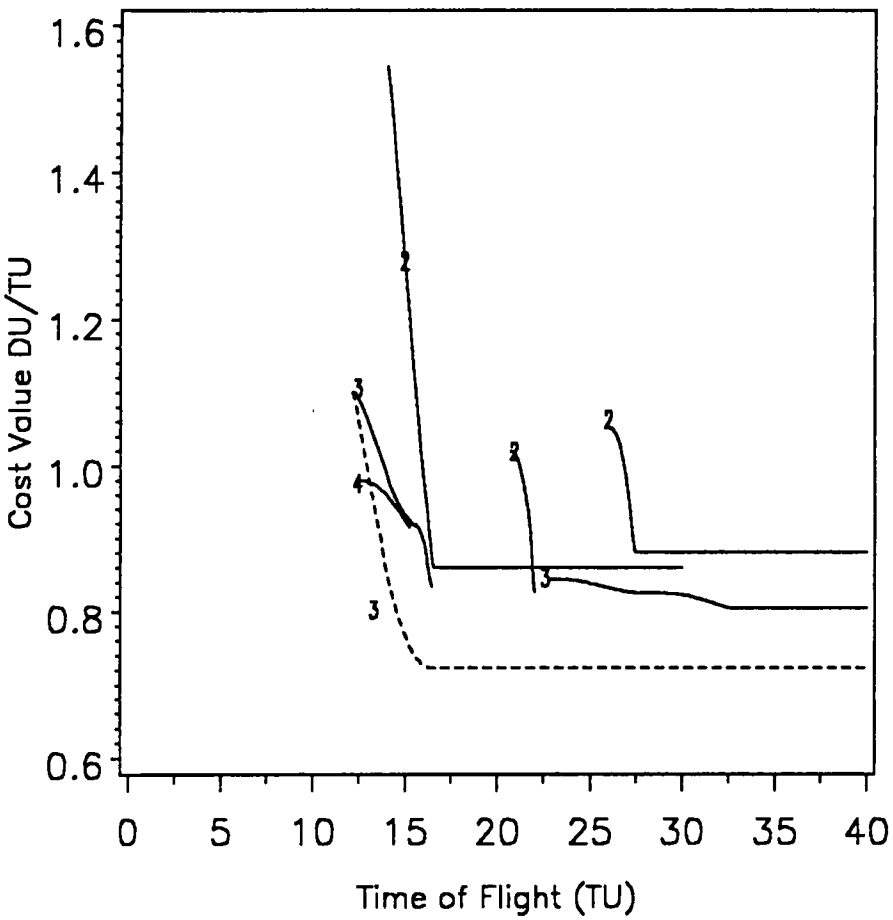

Figure 25. Characteristic velocity versus time to rendezvous for $\beta_T = 180°$, $i_T = 1$ radian.: the minimum radius limit = 1 DU.,

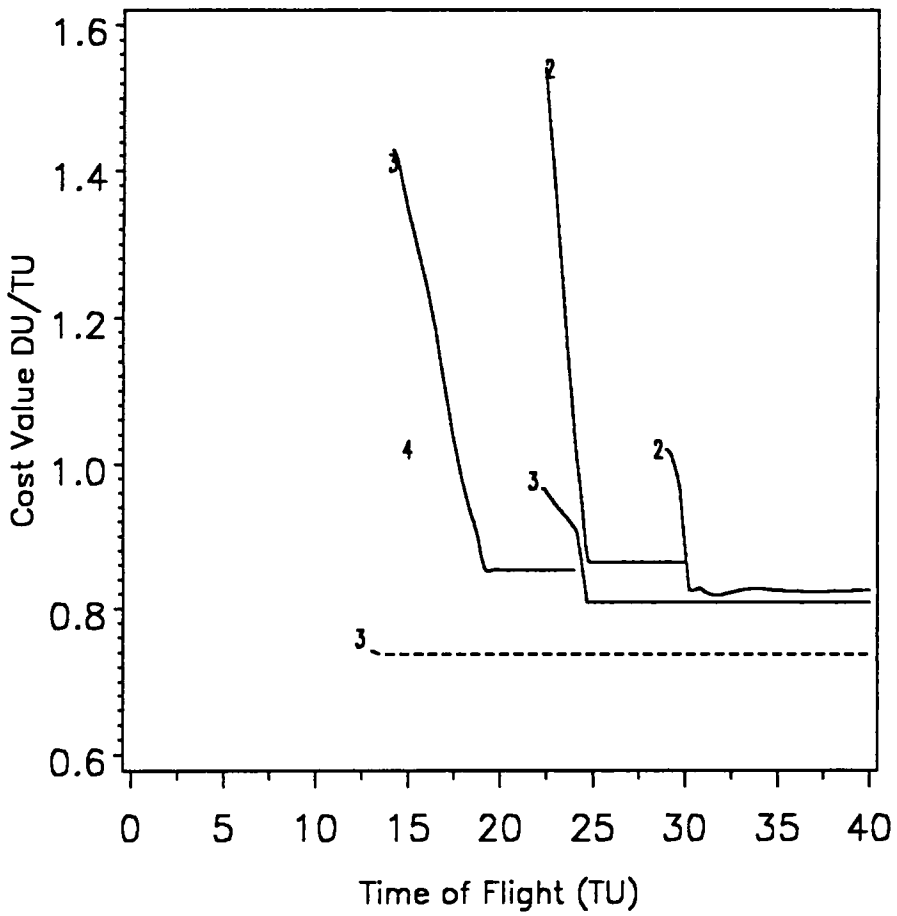

Figure 26. Characteristic velocity versus time to rendezvous for $\beta_T = 270°$, $i_T = 1$ radian.: the minimum radius limit = 1 DU.,

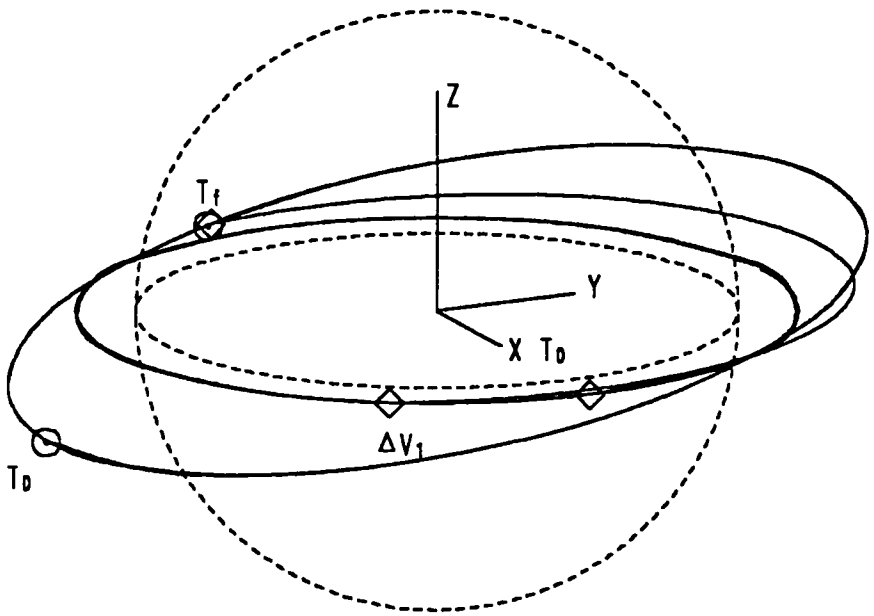

Figure 27. 2 impulse trajectory, $\beta_T = 270°$, $i_T = 0.2$ radians,: cost = 0.2215 DU/TU. Time of flight = 30.35 TU (408.17 minutes).

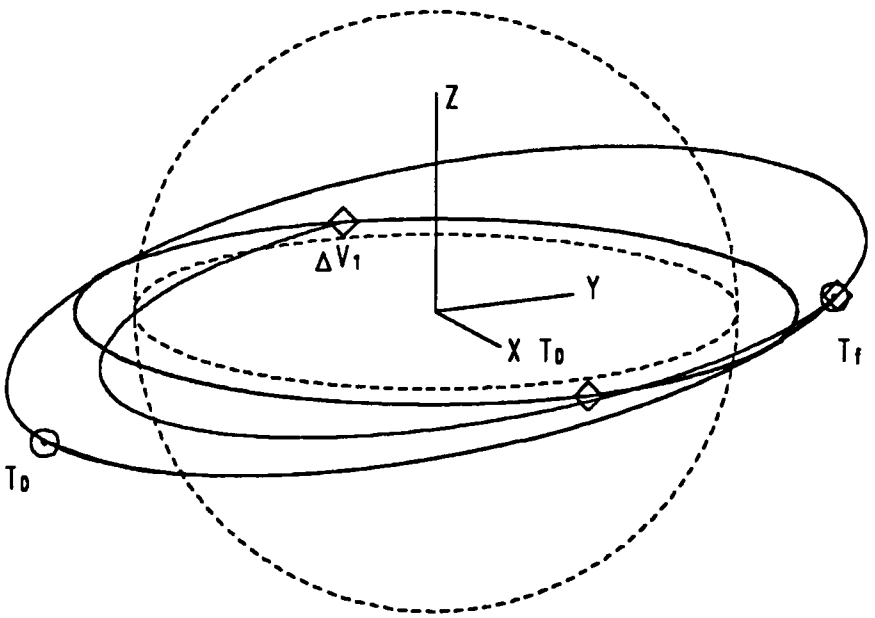

Figure 28. 2 impulse trajectory, $\beta_T = 270°$, $i_T = 0.2$ radians,: cost = 0.2547 DU/TU. Time of flight = 26.01 TU (349.73 minutes).

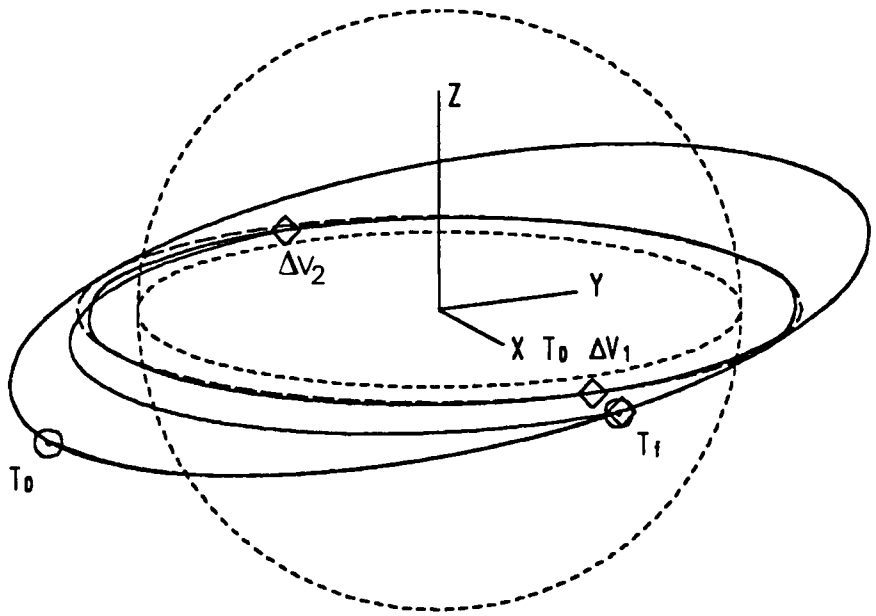

Figure 29. 3 impulse trajectory, $\beta_T = 270°$, $i_T = 0.2$ radians,: cost = 0.1998 DU/TU. Time of flight = 24.67 TU (331.73 minutes).

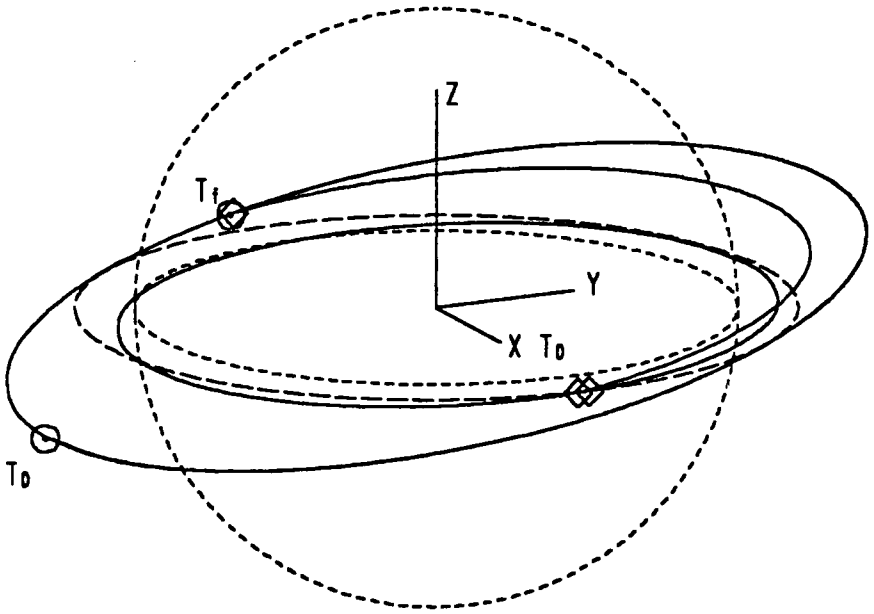

Figure 30. 3 impulse trajectory, $\beta_T = 270°$, $i_T = 0.2$ radians,: cost = 0.2439 DU/TU. Time of flight = 19.28 TU (259.31 minutes).

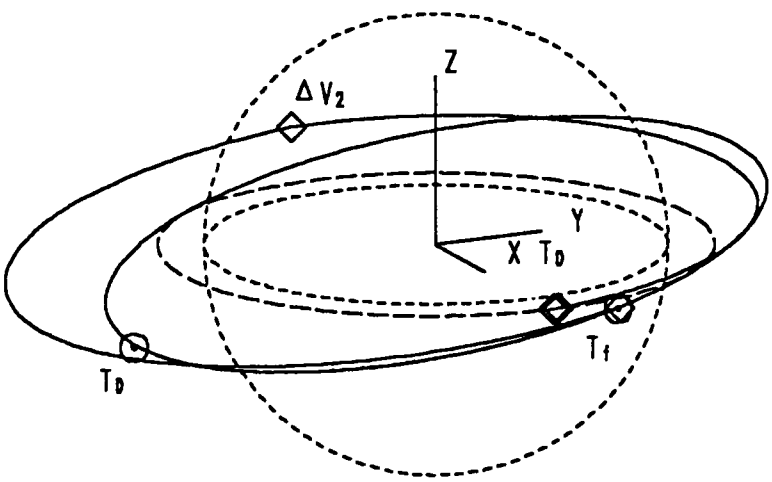

Figure 31. 3 impulse trajectory, $\beta_T = 270°$, $i_T = 0.2$ radians,: cost = 0.2702 DU/TU. Time of flight = 14.29 TU (192.10 minutes).

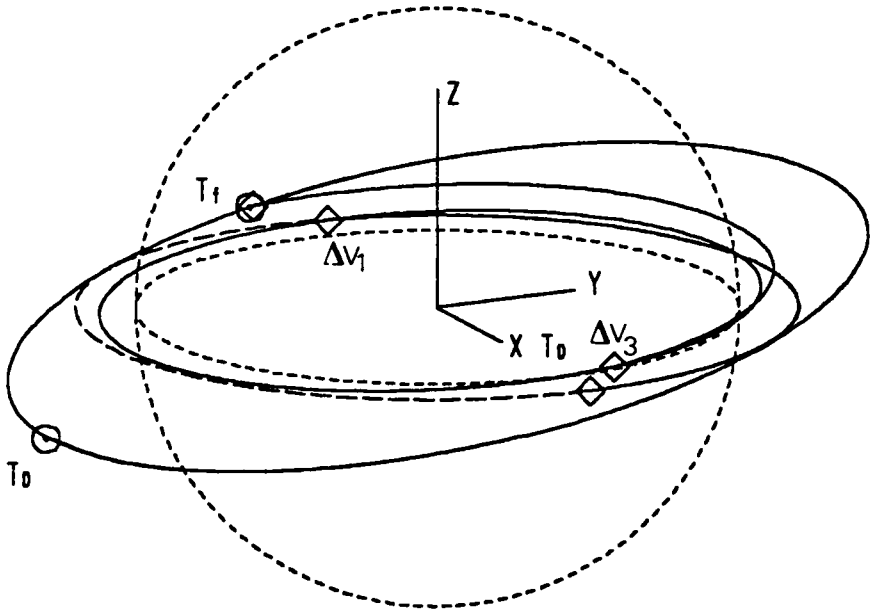

Figure 32. 4 impulse trajectory, $\beta_T = 270°$, $i_T = 0.2$ radians,: cost = 0.2403 DU/TU. Time of flight = 19.19 TU (258.14 minutes).

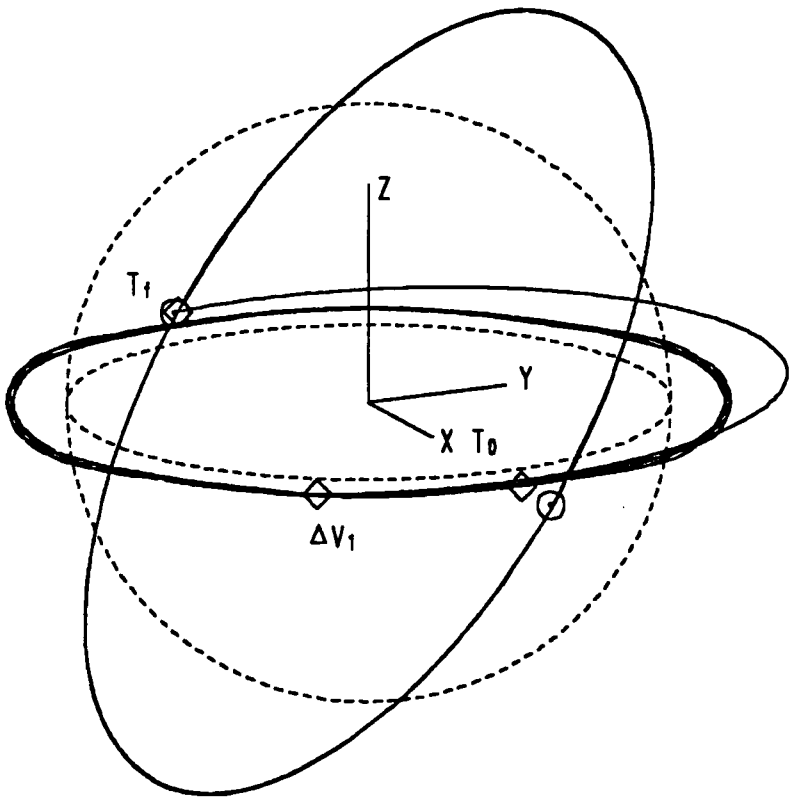

Figure 33. 2 impulse trajectory, $\beta_T = 0°$, $i_T = 1.0$ radians,: cost = 0.8233 DU/TU. Time of flight = 38.45 TU (516.98 minutes).

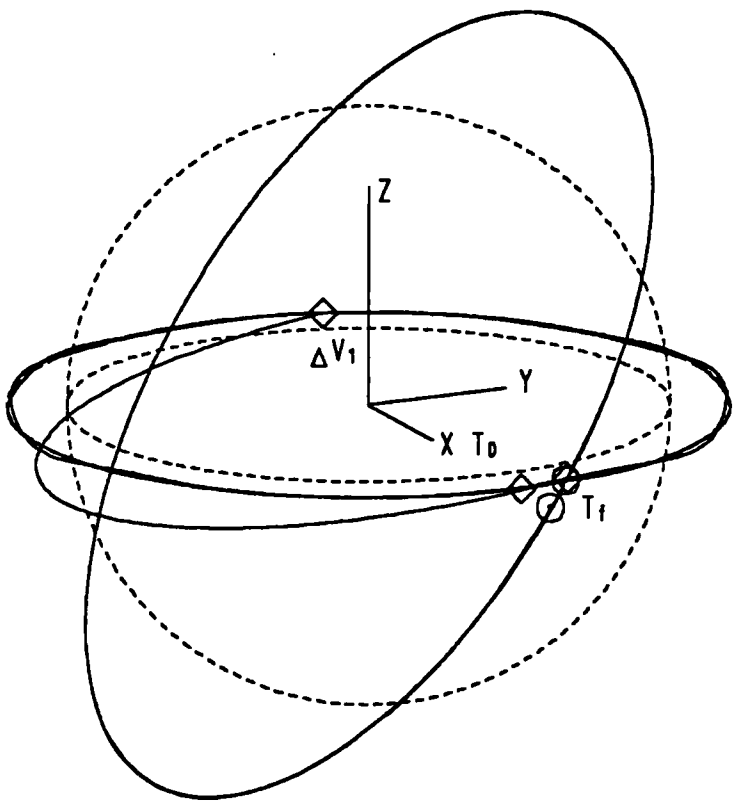

Figure 34. 2 impulse trajectory, $\beta_T = 0°$, $i_T = 1.0$ radians,: cost = 0.8669 DU/TU. Time of flight = 33.04 TU (444.25 minutes).

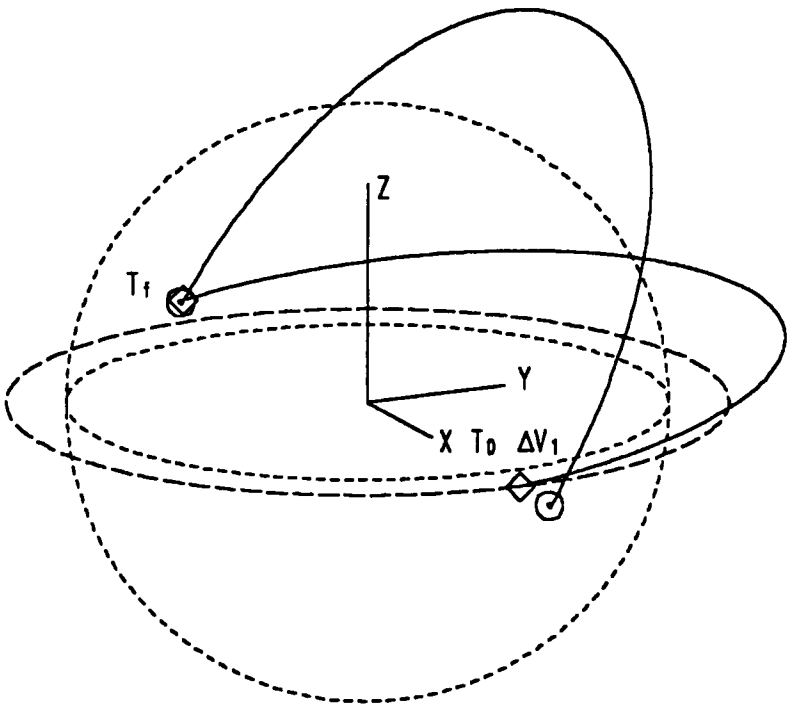

Figure 35. 2 impulse trajectory, $\beta_T = 0°$, $i_T = 1.0$ radians,: cost = 0.8447 DU/TU. Time of flight = 5.49 TU (73.76 minutes).

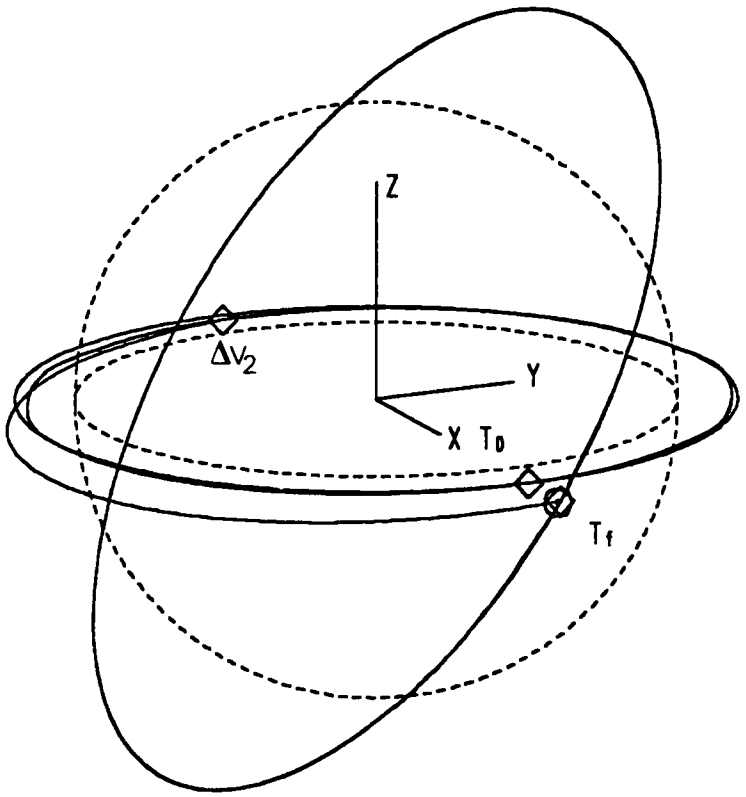

Figure 36. 3 impulse trajectory, $\beta_T = 0°$, $i_T = 1.0$ radians,: cost = 0.8085 DU/TU. Time of flight = 32.91 TU (442.53 minutes).

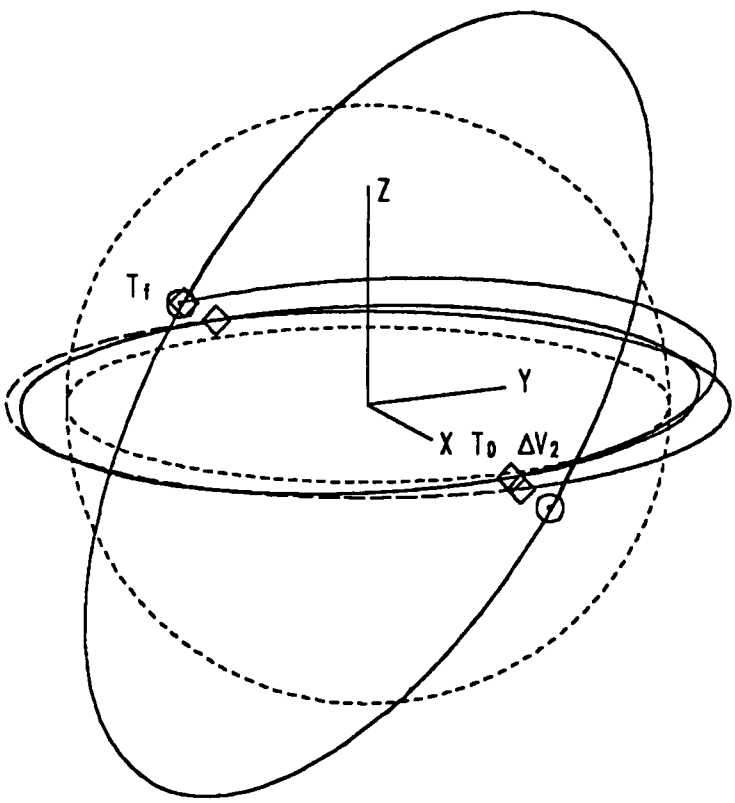

Figure 37. 3 impulse trajectory, $\beta_T = 0°$, $i_T = 1.0$ radians,: cost = 0.8154 DU/TU. Time of flight = 27.42 TU (368.77 minutes).

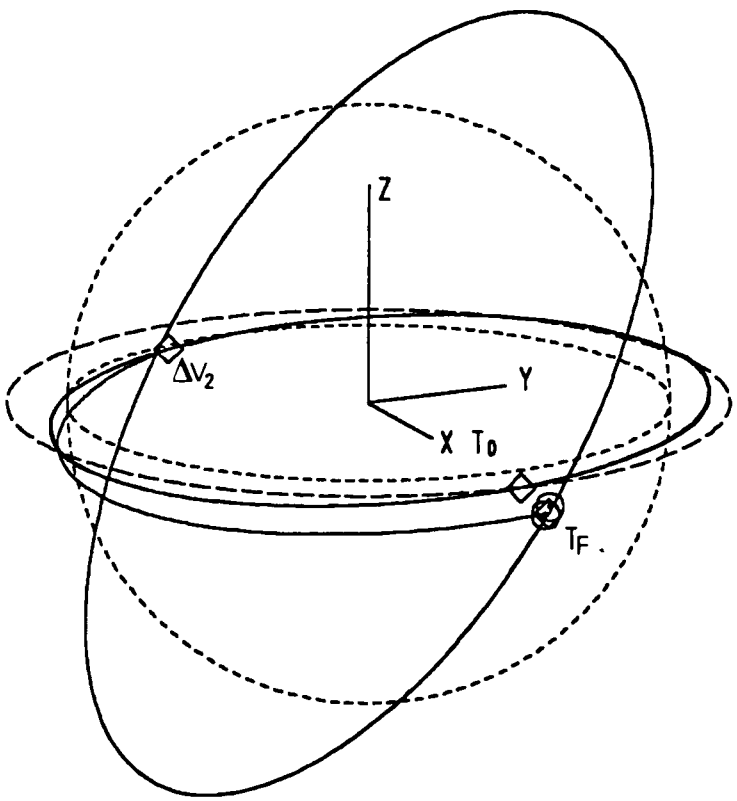

Figure 38. 3 impulse trajectory, $\beta_T = 0°$, $i_T = 1.0$ radians,: cost = 0.8659 DU/TU. Time of flight = 21.90 TU (294.55 minutes). The minimum radius constraint is active (2 touch point).

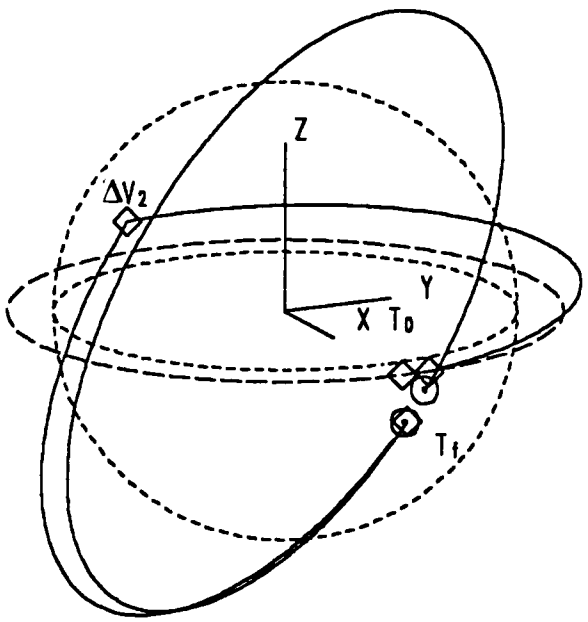

Figure 39. 3 impulse trajectory, $\beta_T = 0°$, $i_T = 1.0$ radians,: cost = 0.7789 DU/TU. Time of flight = 10.77 TU (144.86 minutes).

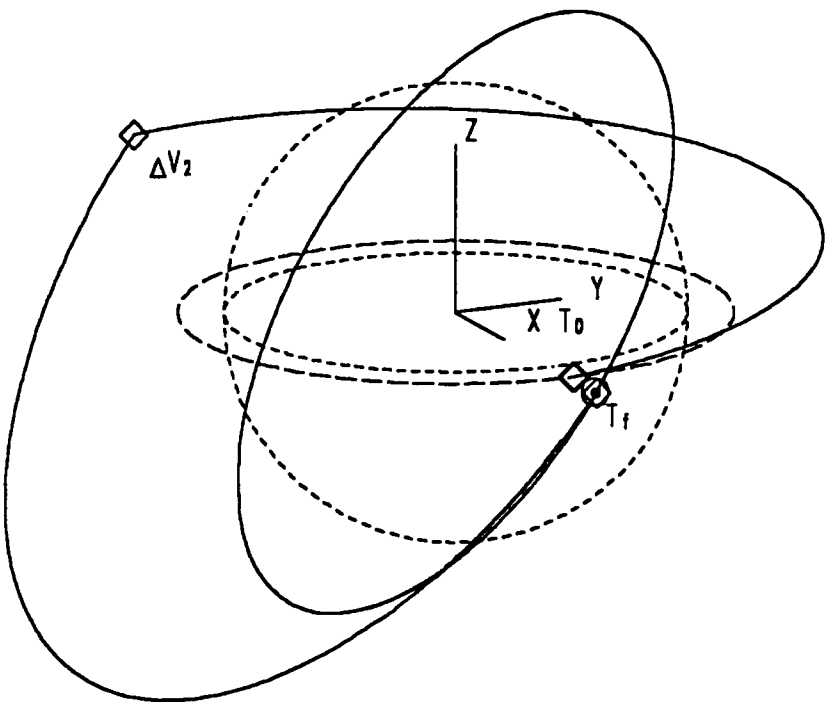

Figure 40. 3 impulse trajectory, $\beta_T = 0°$, $i_T = 1.0$ radians,: cost = 0.7141 DU/TU. Time of flight = 21.93 TU (294.88 minutes).

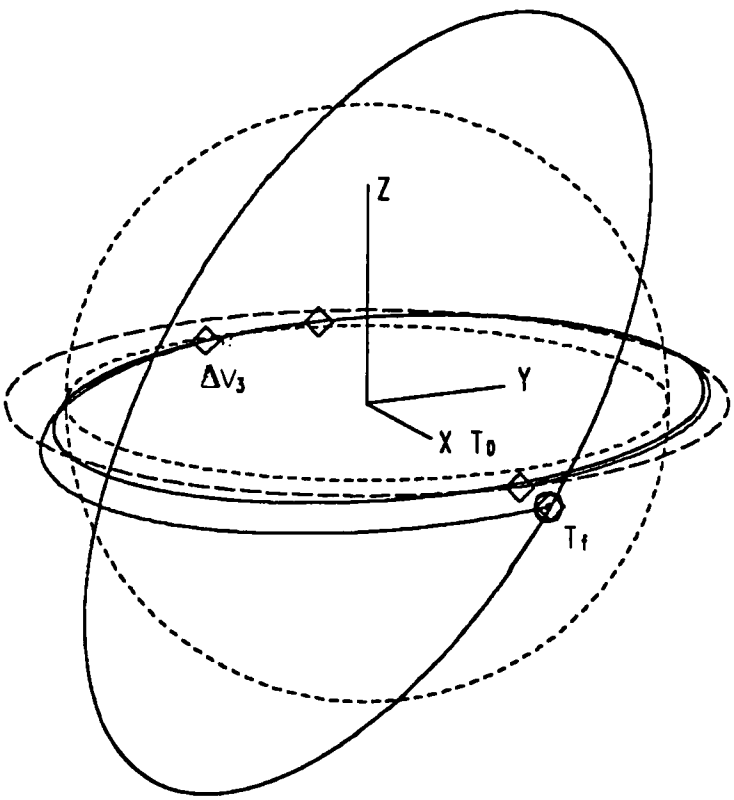

Figure 41. 4 impulse trajectory, $\beta_T = 0°$, $i_T = 1.0$ radians,: cost = 0.8319 DU/TU. Time of flight = 21.93 TU (294.93 minutes). The minimum radius constraint is active (1 touch point on the 3rd arc).

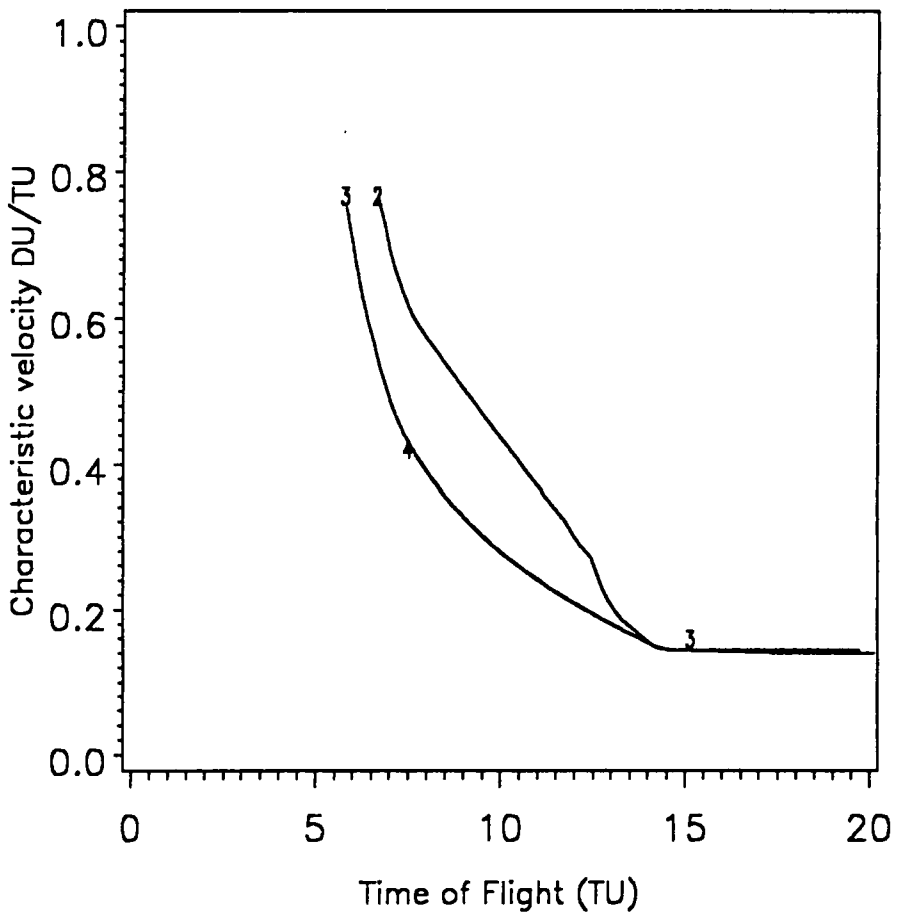

Figure 42. Characteristic velocity versus time to rendezvous for the ellipse-to-ellipse case.: the minimum radius limit = 1 DU., $\beta_T = 120°$, $i_T = 0$ radian.

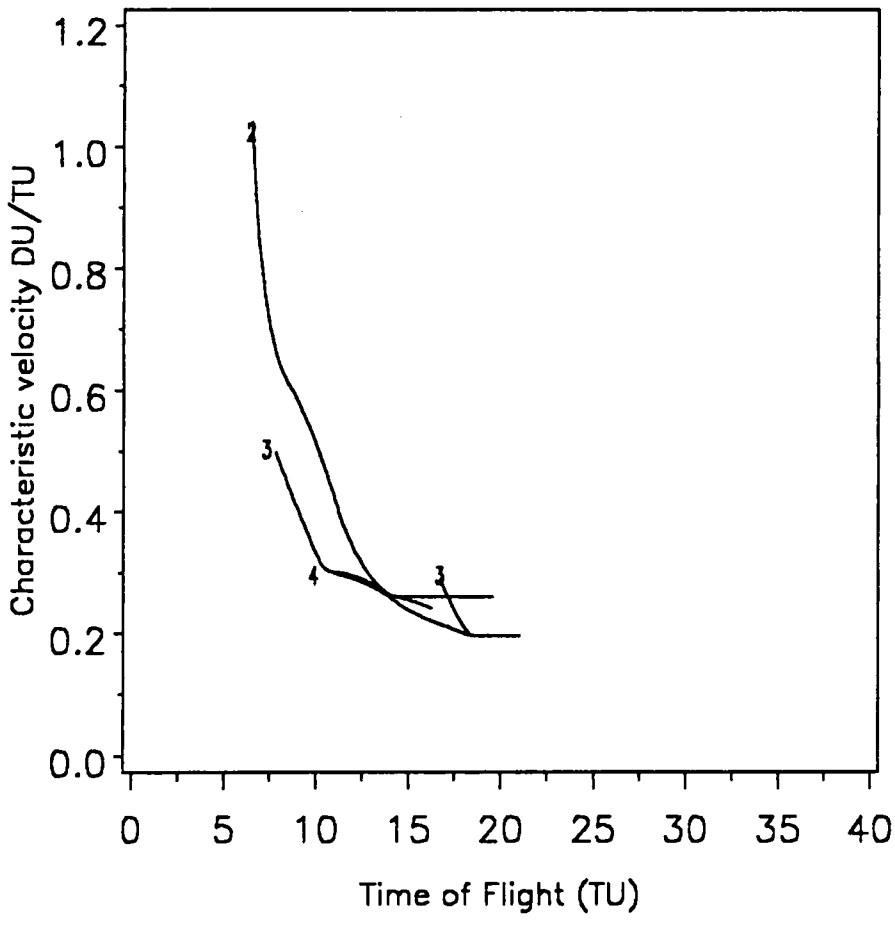

Figure 43. Characteristic velocity versus time to rendezvous for the ellipse-to-ellipse case.: the minimum radius limit = 1 DU., $\beta_T = 120°$, $i_T = 0.2$ radians.

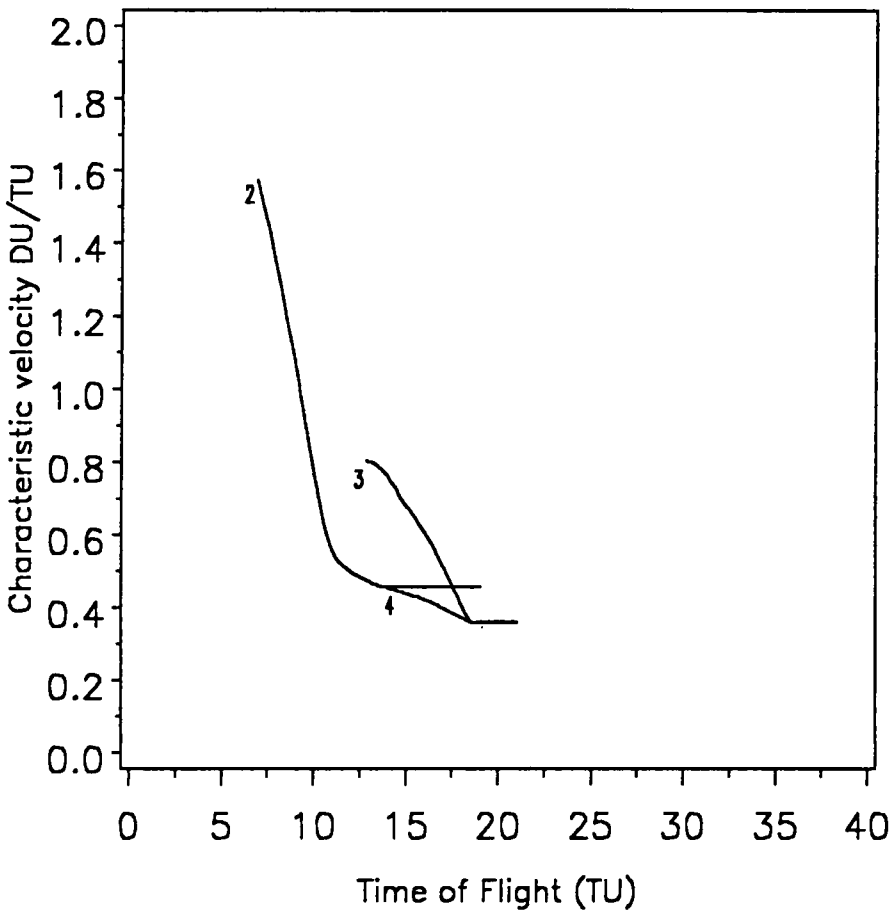

Figure 44. Characteristic velocity versus time to rendezvous for the ellipse-to-ellipse case.: the minimum radius limit = 1 DU., $\beta_T = 120°$, $i_T = 0.5$ radians.

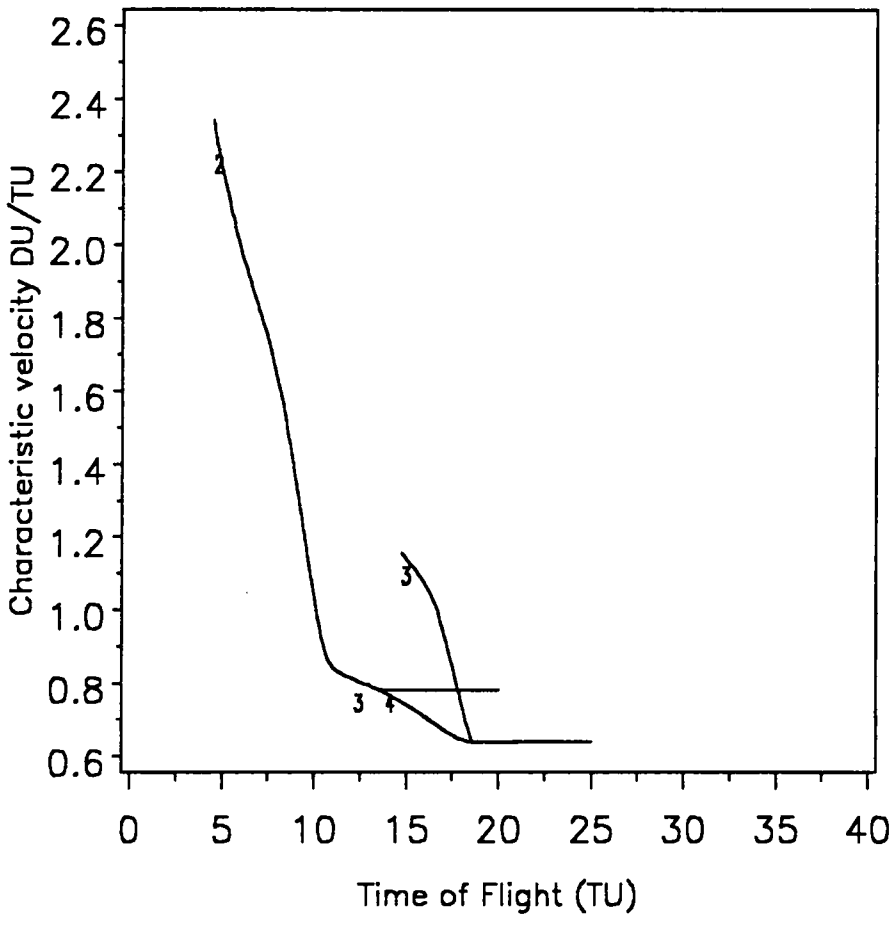

Figure 45. Characteristic velocity versus time to rendezvous for the ellipse-to-ellipse case.: the minimum radius limit = 1 DU., $\beta_T = 120°$, $i_T = 1$ radian.

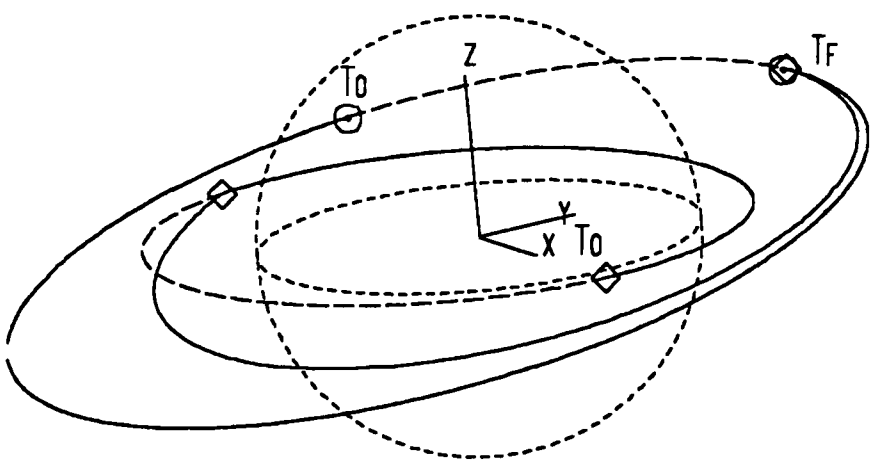

Figure 46. 2 impulse trajectory.: Target characteristics: a = 2.0, e = 0.1, i = 0.2 radians, $\Omega_T = 0°$, $\omega_T = 45°$, $\beta_T = 120°$. Interceptor characteristics: a = 1.4, e = 0.2, i = 0.0, $\Omega_i = 0°$, $\omega_i = 0°$. Cost = 0.2599 DU/TU. Time of flight = 14.57 TU (195.88 minutes).

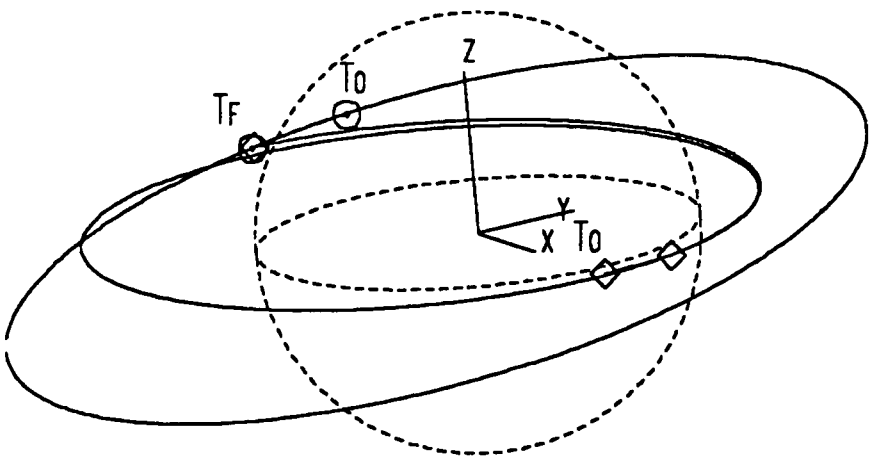

Figure 47. 3 impulse trajectory.: Target characteristics: a = 2.0, e = 0.1, i = 0.2 radians, $\Omega_T = 0°$, $\omega_T = 45°$, $\beta_T = 120°$. Interceptor characteristics: a = 1.4, e = 0.2, i = 0.0, $\Omega_i = 0°$, $\omega_i = 0°$. Cost = 0.1969 DU/TU. Time of flight = 18.47 TU (248.30 minutes).

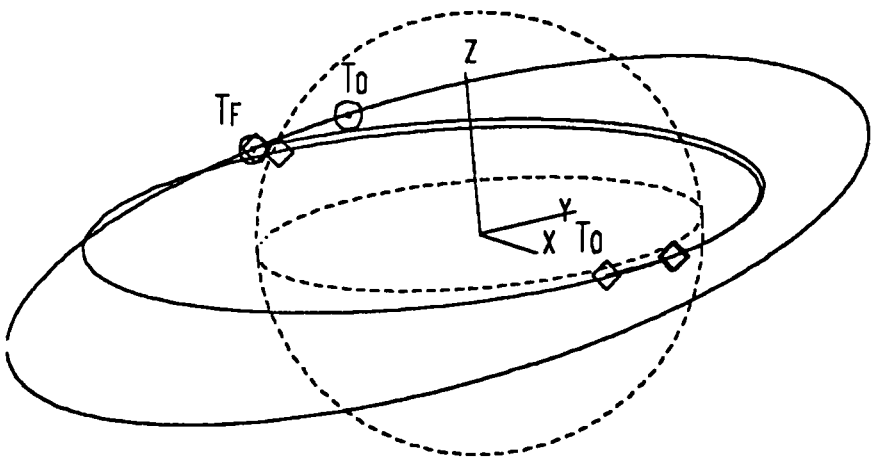

Figure 48. 4 impulse trajectory.: Target characteristics: a = 2.0, e = 0.1, i = 0.2 radians, $\Omega_T = 0°$, $\omega_T = 45°$, $\beta_T = 120°$. Interceptor characteristics: a = 1.4, e = 0.2, i = 0.0, $\Omega_i = 0°$, $\omega_i = 0°$. Cost = 0.1972 DU/TU. Time of flight = 18.47 TU (248.32 minutes).

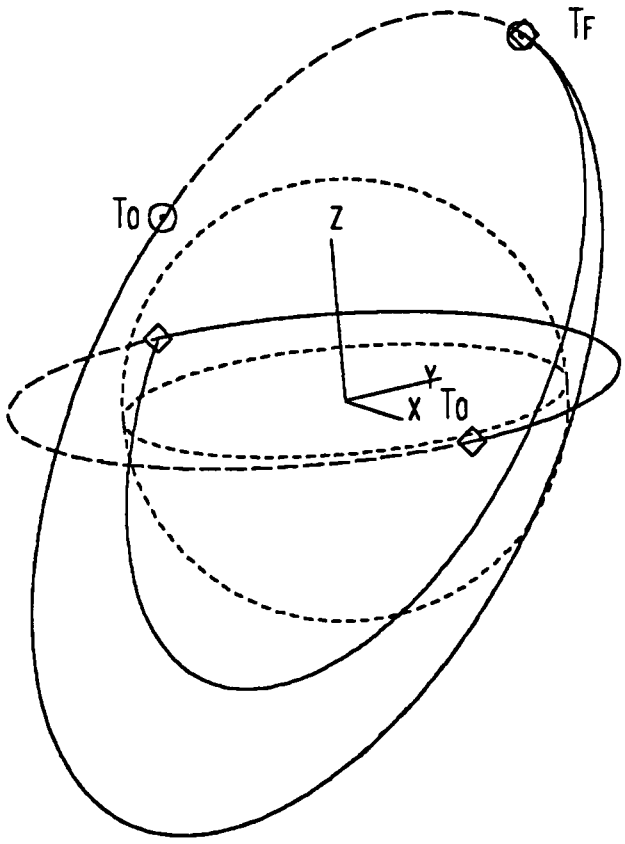

Figure 49. 2 impulse trajectory.: Target characteristics: a = 2.0, e = 0.1, i = 1.0, $\Omega_T = 0°$, $\omega_T = 45°$, $\beta_T = 120°$. Interceptor characteristics: a = 1.4, e = 0.2, i = 0.0, $\Omega_i = 0°$, $\omega_i = 0°$. Cost = 0.7796 DU/TU. Time of flight = 13.98 TU (188.04 minutes).

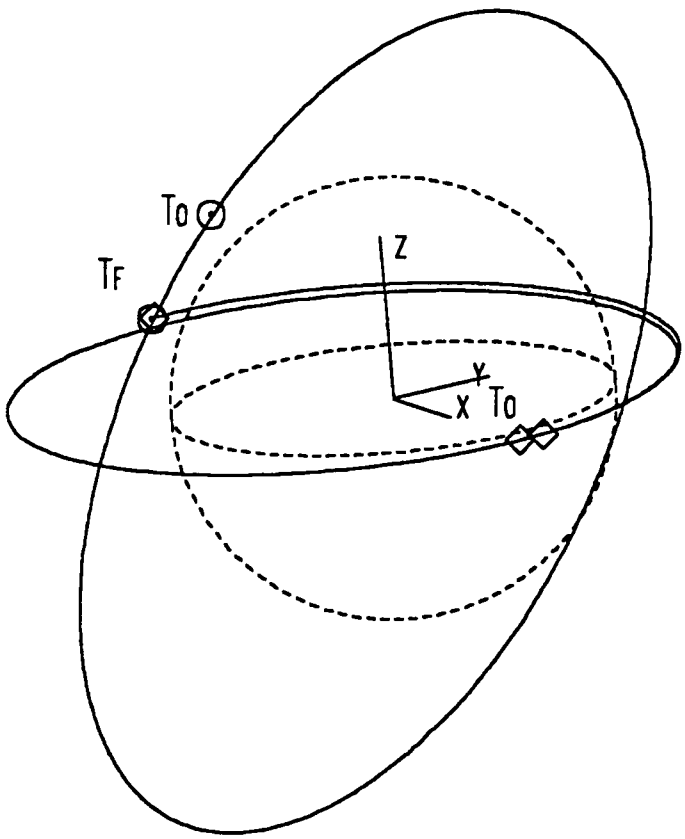

Figure 50. 3 impulse trajectory.: Target characteristics: a = 2.0, e = 0.1, i = 1.0, $\Omega_T = 0°$, $\omega_T = 45°$, $\beta_T = 120°$. Interceptor characteristics: a = 1.4, e = 0.2, i = 0.0, $\Omega_i = 0°$, $\omega_i = 0°$. Cost = 0.6369 DU/TU. Time of flight = 18.59 TU (249.93 minutes).

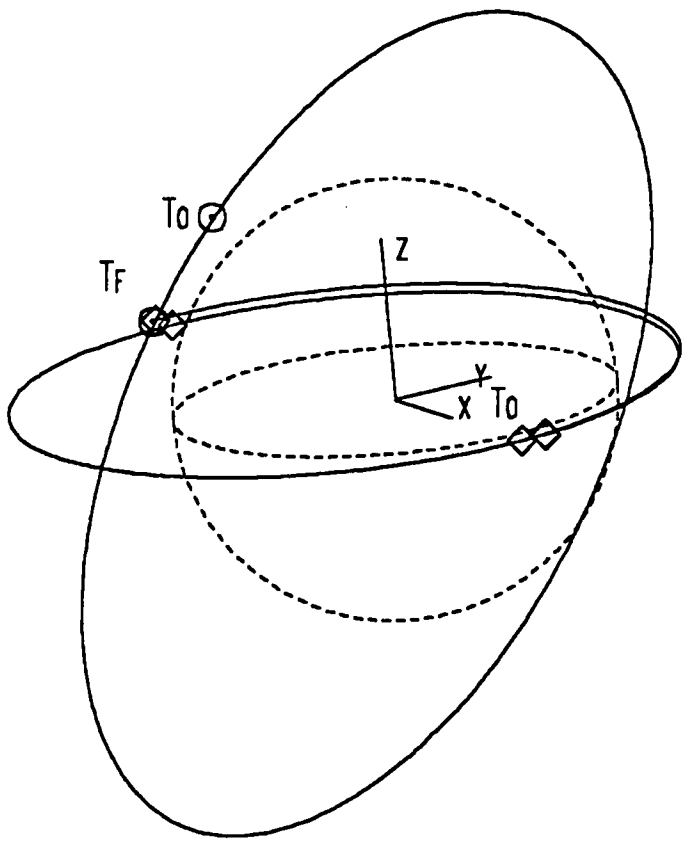

Figure 51. 4 impulse trajectory. Target characteristics: a = 2.0, e = 0.1, i = 1.0, $\Omega_T = 0°, \omega_T = 45°, \beta_T = 120°$. Interceptor characteristics: a = 1.4, e = 0.2, i = 0.0, $\Omega_i = 0°, \omega_i = 0°$. Cost = 0.6372 DU/TU. Time of flight = 18.59 TU (249.93 minutes).

Appendix A. Time of flight derivatives.

Time of flight of the interceptor can be written as:

$$T_I = \sum_{i=1}^{nim} \Delta t_i, \qquad [A.1]$$

where Δt_i is the transfer time on the i^{th} subarc and is given by

$$\Delta t_i = \int_0^{\eta_i} \frac{d\eta}{h_i u_i^2}. \qquad [A.2]$$

Therefore, the derivative with respect to any of the variables is best obtained using the chain rule. Or

$$\frac{\partial T_I}{\partial \eta_j} = \frac{1}{h_j u_{j+1}^2} + \sum_{i=j+1}^{nim} \left\{ \frac{\partial \Delta t_i}{\partial u_i} \frac{\partial u_i}{\partial \eta_j} + \frac{\partial \Delta t_i}{\partial u'_i} \frac{\partial u'_i}{\partial \eta_j} \right\}, \qquad [A.3]$$

$$\frac{\partial T_I}{\partial \Delta u'_j} = \sum_{i=j+1}^{nim} \left\{ \frac{\partial \Delta t_i}{\partial u_i} \frac{\partial u_i}{\partial \Delta u'_j} + \frac{\partial \Delta t_i}{\partial u'_i} \frac{\partial u'_i}{\partial \Delta u'_j} \right\}, \qquad [A.4]$$

$$\frac{\partial T_I}{\partial \Delta h_j} = \sum_{i=j+1}^{nim} \left\{ \frac{\partial \Delta t_i}{\partial u_i} \frac{\partial u_i}{\partial \Delta u'_j} + \frac{\partial \Delta t_i}{\partial u'_i} \frac{\partial u'_i}{\partial \Delta h_j} + \frac{\partial \Delta t_i}{\partial h_i} \frac{\partial h_i}{\partial \Delta h_j} \right\}, \qquad [A.5]$$

The partial derivatives $\{ \frac{\partial \Delta t_i}{\partial u_i(0)}, \frac{\partial \Delta t_i}{\partial u'_i(0)}, \frac{\partial \Delta t_i}{\partial h_i} \}$ for each subarc is obtained for each type of orbit, i.e., elliptic (but non-circular), parabolic, and hyperbolic as:

$$\frac{\partial \Delta t_i}{\partial u_i(0)} = \frac{-2 h_i^5}{\mu^3} \left\{ \frac{\sin(v_0)}{2 e} \left[\frac{1}{[1 + e \cos(f)]^2} \right]_{v_0}^{v_f} + \cos(v_0) I_1 \right\}, \qquad [A.6]$$

$$\frac{\partial \Delta t_i}{\partial u'_i(0)} = \frac{-2 h_i^5}{\mu^3} \left\{ \frac{\cos(v_0)}{2 e} \left[\frac{1}{[1 + e \cos(f)]^2} \right]_{v_0}^{v_f} - \sin(v_0) I_1 \right\}, \qquad [A.7]$$

$$\frac{\partial \Delta t_i}{\partial h_i} = \frac{4 h_i^2}{\mu^2} \left\{ I_2 - \frac{\sin(v_0)}{2 e} \left[\frac{1}{1 + e \cos(f)} \right]_{v_0}^{v_f} - \cos(v_0) I_1 \right\} - \frac{\Delta t_i}{h_i}, \qquad [A.8]$$

where

$$I_1 = \int_{v_0}^{v_f} \frac{\cos(f) df}{[1 + e \cos(f)]^3}. \qquad [A.9]$$

For $e < 1$, this is given by

$$I_1 = \frac{(1+e^2)}{(1-e^2)^2}\left[\frac{\sin(f)}{1+e\cos(f)}\right]_{v_0}^{v_f} - \frac{3\,e\,E}{2\,(1-e^2)^{5/2}}\Bigg|_{E(v_0)}^{E(v_f)} - \frac{e}{2\,(1-e^2)^2}\left\{\frac{\sin(f)\,(\cos(f)+e)}{[1+e\cos(f)]^2}\right\}_{v_0}^{v_f}, \qquad [A.10]$$

where $E(v)$ is the eccentric anomaly and is obtained from

$$\sin(E) = \frac{\sin(v)}{au\sqrt{1-e^2}} \quad \text{and} \qquad \cos(E) = \frac{e+\cos(v)}{1+e\cos(v)}. \qquad [A.12]$$

for $e = 1$.

$$I_1 = \left[\frac{1}{8}\tan(\frac{f}{2}) - \frac{1}{40}\tan^5(\frac{f}{2})\right]_{v_0}^{v_f}, \qquad [A.12]$$

for $e > 1$,

$$I_1 = \frac{(1+e^2)}{(e^2-1)^2}\left[\frac{\sin(f)}{1+e\cos(f)}\right]_{v_0}^{v_f} - \frac{3\,e\,F}{2\,(e^2-1)^{5/2}}\Bigg|_{F(v_0)}^{F(v_f)} - \frac{e}{2\,(e^2-1)^2}\left[\frac{\sin(f)\,(\cos(f)+e)}{[1+e\cos(f)]^2}\right]_{v_0}^{v_f}, \qquad [A.13]$$

where $F(v)$ is the hyperbolic eccentric anomaly and is obtained from

$$\text{Let } y = \frac{e+\cos(v)}{1+e\cos(v)}, \quad \text{then} \qquad F = \ln[y + \sqrt{y^2-1}\,]. \qquad [A.15]$$

$$I_2 = \int_{v_0}^{v_f} \frac{df}{[1 + e\cos(f)]^3}. \qquad [A.14]$$

For $e < 1$

$$I_2 = \frac{e^2}{2(1-e^2)^2}\left[\frac{\sin(f)[e + \cos(f)]}{2(1-e^2)^2[1 + e\cos(f)]^2}\right]_{v_0}^{v_f} + \frac{(2+e^2)E}{2(1-e^2)^{5/2}}\Bigg|_{E(v_0)}^{E(v_f)} - \frac{2e\sin(f)}{(1-e^2)^2[1+e\cos(f)]}\Bigg|_{v_0}^{v_f}, \qquad [A.15]$$

for $e = 1$

$$I = \frac{z}{4}\left[1 + \frac{2z^2}{3} + \frac{z^4}{5}\right]_{\tan(\frac{v_0}{2})}^{\tan(\frac{v_f}{2})}, \qquad [A.16]$$

and for $e > 1$

$$I_2 = \frac{e^2}{2(e^2-1)^2}\left[\frac{\sin(f)(e + \cos(f))}{2(e^2-1)^2(1 + e\cos(f))^2}\right]_{v_0}^{v_f} + \frac{(2+e^2)F}{2(e^2-1)^{5/2}}\Bigg|_{F(v_0)}^{F(v_f)} - \frac{2e\sin(f)}{(e^2-1)^2(1+e\cos(f))}\Bigg|_{v_0}^{v_f}, \qquad [A.17]$$

where E, F are the eccentric anomaly on an elliptic orbit and a hyperbolic orbit respectively, v_0 is the true anomaly of the intial point on each subarc, η_i is the transfer angle on each subarc and $v_f = v_0 + \eta_i$.

For a near circular subarc, i.e., if e is very small, the derivatives with respect to the initial conditions on that subarc are given by

$$\frac{\partial \Delta t_i}{\partial u_i(0)} = \frac{3eh_i^5}{\mu^3}\left\{\cos(v_0)\left[\eta_i + \sin(\eta_i)\cos(\eta_i)\right] - \sin(v_0)\,\sin^2(\eta_i)\right\} - \frac{2h_i^5}{\mu^3}\sin(\eta_i) \qquad [A.18]$$

$$\frac{\partial \Delta t_i}{\partial u'_i(0)} = \frac{3eh_i^5}{\mu^3}\left\{\cos(v_0)\,\sin^2(\eta_i) - \sin(v_0)\left[\eta_i - \sin(\eta_i)\cos(\eta_i)\right]\right\} + \frac{2h_i^5}{\mu^3}\left[\cos(\eta_i) - 1\right] \qquad [A.19]$$

$$\frac{\partial \Delta t_i}{\partial h_i} = \frac{-6eh_i^2}{\mu^2}\left\{2\left(\sin(v_f) - \sin(v_0)\right) - \cos(v_0)\left[\eta_i + \sin(\eta_i)\cos(\eta_i)\right] + \sin(v_0)\,\sin^2(\eta_i)\right\} - \frac{4h_i^2}{\mu^2}\left[\eta_i - \sin(\eta_i)\right] - \frac{\Delta t_i}{h_i} \qquad [A.20]$$

Second derivatives of the time of flight expressions may be obtained in a like manner.

Appendix B. Derivatives of the initial conditions on a subarc.

The derivatives of the initial conditions on the m^{th} subarc with respect to the variables $\{\eta, \Delta u', \Delta h, \phi\}_l$ are given as follows:

i. the derivatives of $\{u_m(0), u'_m(0), h_m\}$ with respect to η_l. Let $\zeta = \sum_{i=l}^{m-1} \eta_i$.

$$\frac{\partial u_m(0)}{\partial \eta_l} = -\left[u_l(0) - \frac{\mu}{h_l^2}\right] \sin(\zeta) + u'_l(0) \cos(\zeta), \qquad [B.1]$$

$$\frac{\partial u'_m(0)}{\partial \eta_l} = -\left[u_l(0) - \frac{\mu}{h_l^2}\right] \cos(\zeta) - u'_l(0) \sin(\zeta). \qquad [B.2]$$

$$\frac{\partial h_m}{\partial \eta_l} = 0 \qquad [B.3]$$

If $l \geq m$ then the above derivatives are zero.

ii. the derivatives of $\{u_m(0), u'_m(0), h_m\}$ with respect to $\Delta u'_l$.

If $l < m-1$ then let $\zeta = \sum_{i=l+1}^{m-1} \eta_i$.

$$\frac{\partial u_m}{\partial \Delta u'_l} = \sin(\zeta), \qquad [B.4]$$

$$\frac{\partial u'_m}{\partial \Delta u'_l} = \cos(\zeta). \qquad [B.5]$$

If $l = m-1$ then

$$\frac{\partial u_m}{\partial \Delta u'_l} = 0, \quad \frac{\partial u'_m}{\partial \Delta u'_l} = 1. \qquad [B.6]$$

$$\frac{\partial h_m}{\partial \Delta u'_l} = 0, \quad \text{for all } l. \qquad [B.7]$$

iii. the derivatives of $\{u_m(0), u'_m(0), h_m\}$ with respect to Δh_l.

If $l < m-1$ then the partial derivatives of $u_m(0)$ and $u'_m(0)$ are best expressed in the form of a recursive equation. Let $\{a^1, b^1\} = 0$, then

$$\begin{bmatrix} \dfrac{\partial u_m(0)}{\partial \Delta h_l} \\ \dfrac{\partial u'_m(0)}{\partial \Delta h_l} \end{bmatrix} = \begin{bmatrix} a^{m-l} \\ b^{m-l} \end{bmatrix}. \qquad [B.8]$$

where $[a^k, b^k]$ is given by the recursive equation

$$\begin{bmatrix} a^{k+1} \\ b^{k+1} \end{bmatrix} = \begin{bmatrix} \cos(\eta_{k+l-1}) & \sin(\eta_{k+l-1}) \\ -\sin(\eta_{k+l-1}) & \cos(\eta_{k+l-1}) \end{bmatrix} \begin{bmatrix} a^k \\ b^k \end{bmatrix} - \frac{2\mu}{h^3_{k+l-1}} \begin{bmatrix} 1-\cos(\eta_{k+l-1}) \\ \sin(\eta_{k+l-1}) \end{bmatrix}, \qquad [B.9]$$

for $k = 1, \ldots, m-l-1$.

If $l < m$ then

$$\frac{\partial h_m}{\partial \Delta h_l} = 1. \qquad [B.10]$$

For all other cases and with respect to ϕ_l the partial derivatives of $\{u_m, u', h_m\}$ are zero.

To obtain the derivatives of the unit vector quantities, we note that they are dependent only on the two angle variables η and ϕ . Thus the initial unit vector quantities on the mth subarc are obtained as follows: If

$$[P]_l^m = \prod_{i=l}^{m} \left[\begin{bmatrix} \mathrm{I} & 0 & 0 \\ 0 & \cos(\phi_i)\,\mathrm{I} & \sin(\phi_i)\,\mathrm{I} \\ 0 & -\sin(\phi_i)\,\mathrm{I} & \cos(\phi_i)\,\mathrm{I} \end{bmatrix} \begin{bmatrix} \cos(\eta_i)\,\mathrm{I} & \sin(\eta_i)\,\mathrm{I} & 0 \\ -\sin(\eta_i)\,\mathrm{I} & \cos(\eta_i)\,\mathrm{I} & 0 \\ 0 & 0 & \mathrm{I} \end{bmatrix} \right], \qquad [B.11]$$

then

$$\begin{bmatrix} \hat{r}_m(0) \\ \hat{r}'_m(0) \\ \hat{h}_m(0) \end{bmatrix} = [P]_1^m \begin{bmatrix} \hat{r}_1(0) \\ \hat{r}'_1(0) \\ \hat{h}_1(0) \end{bmatrix}. \qquad [B.12]$$

Thus the derivatives with respect to η_l are given by

$$\begin{bmatrix} \frac{\partial \hat{r}_m(0)}{\partial \eta_l} \\ \frac{\partial \hat{r}'_m(0)}{\partial \eta_l} \\ \frac{\partial \hat{h}_m(0)}{\partial \eta_l} \end{bmatrix} = [P]_{l+1}^{m} \begin{bmatrix} \mathrm{I} & 0 & 0 \\ 0 & \cos(\phi_l)\,\mathrm{I} & \sin(\phi_l)\,\mathrm{I} \\ 0 & -\sin(\phi_l)\,\mathrm{I} & \cos(\phi_l)\,\mathrm{I} \end{bmatrix} \begin{bmatrix} -\sin(\eta_i)\,\mathrm{I} & \cos(\eta_i)\,\mathrm{I} & 0 \\ -\cos(\eta_i)\,\mathrm{I} & -\sin(\eta_i)\,\mathrm{I} & 0 \\ 0 & 0 & 0 \end{bmatrix} [P]_1^{l-1} \begin{bmatrix} \hat{r}_1(0) \\ \hat{r}'_1(0) \\ \hat{h}_1(0) \end{bmatrix}. \qquad [B.13]$$

The derivatives with respect to ϕ_l are given by

$$\begin{bmatrix} \frac{\partial \hat{r}_m(0)}{\partial \phi_l} \\ \frac{\partial \hat{r}'_m(0)}{\partial \phi_l} \\ \frac{\partial \hat{h}_m(0)}{\partial \phi_l} \end{bmatrix} = [P]_{l+1}^{m} \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\sin(\phi_l)\,\mathrm{I} & \cos(\phi_l)\,\mathrm{I} \\ 0 & -\cos(\phi_l)\,\mathrm{I} & -\sin(\phi_l)\,\mathrm{I} \end{bmatrix} \begin{bmatrix} \cos(\eta_l)\,\mathrm{I} & \sin(\eta_l)\,\mathrm{I} & 0 \\ -\sin(\eta_l)\,\mathrm{I} & \cos(\eta_l)\,\mathrm{I} & 0 \\ 0 & 0 & \mathrm{I} \end{bmatrix} [P]_1^{l-1} \begin{bmatrix} \hat{r}_1(0) \\ \hat{r}'_1(0) \\ \hat{h}_1(0) \end{bmatrix}. \qquad [B.14]$$

The partial derivatives with respect to $\Delta u'$ and Δh are obviously zero.

The vita has been removed from the scanned document