CSDL-T-813

AN IMPROVEMENT OF GAUSS' METHOD FOR SOLVING LAMBERT'S PROBLEM

by

ROBIN M. VAUGHAN

B. S., Tulane University
(1981)

Submitted to the Department of
Aeronautics and Astronautics
in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE IN AERONAUTICS AND ASTRONAUTICS

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1983

Signature of Author: _____
Robin M. Vaughan
Department of Aeronautics and Astronautics, 6 May 1983

Certified by: _____
Dr. Richard H. Battin, Adjunct Professor
Thesis Supervisor

Accepted by: _____
Prof. Harold Y. Wachman
Chairman, Departmental Graduate Committee

AN IMPROVEMENT OF GAUSS' METHOD FOR SOLVING LAMBERT'S PROBLEM

by

ROBIN M. VAUGHAN

Submitted to the Department of Aeronautics and Astronautics
on May 6, 1983 in partial fulfillment of the
requirements for the Degree of Master of Science in
Aeronautics and Astronautics

Abstract

The method of solving the two-point, time-constrained, orbital boundary value problem, commonly called Lambert's problem, derived by Gauss in 1801 is today considered the classic solution. A new method closely paralleling Gauss' original one is developed here. This method retains the analytic simplicity of Gauss' method but is functionally superior to it. The derivation of this method is based on a geometric transformation of the problem allowed by Lambert's theorem. This permits the equations to be written using the basic form of Kepler's equation. With this formulation, a free parameter can be introduced and chosen so that the resulting successive substitution process converges rapidly. The final method has almost constant convergence properties for a wide range of problems. A detailed derivation of the new method as well as a discussion of the various steps in the iteration process are presented here. A thorough comparison of Gauss' method and the new method is given which highlights the near uniform behavior of the new method.

Thesis Supervisor: Dr. Richard H. Battin
Title: Adjunct Professor of Aeronautics and Astronautics

ACKNOWLEDGEMENTS

I wish to thank my thesis advisor, Dr. Richard Battin, who originally suggested the method for solving Lambert's problem presented in this thesis. His patience, guidance, and advice through all phases of this work is greatly appreciated. I am also grateful to the technical staff of the Charles Stark Draper Laboratory, especially to my supervisor Edward Bergmann, for their cooperation during this work. Special thanks to Lisa Kern for typing the tables and to Dave Hauger and Mario Santarelli of the Flight Simulation Laboratory at CSDL for providing the computer facilties used to produce this document. My deepest gratitude to my family and friends for their encouragement and support.

This report was prepared under contract NAS9-16023 with the National Aeronautics and Space Administration.

Publication of this report does not constitute approval by the Charles Stark Draper Laboratory or NASA of the findings or conclusions contained herein. It is published solely for the exchange of ideas.

TABLE OF CONTENTS

Chapter 1

Introduction

A PROBLEM that has long been of interest in astrodynamics is the determination of the two-body orbit having a specified time of flight which connects two position vectors. The geometry of this problem for the case of an elliptic orbit is shown in Fig. 1. The vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ are the initial and final positions and the angle θ between them is called the transfer angle. The straight line connecting the termini of $\mathbf{r}_1$ and $\mathbf{r}_2$ is called the chord and is of length c. This problem was first addressed by the Swiss mathematician Leonard Euler. His work, which was published in 1743, developed equations to solve the problem for the case of a parabolic orbit. Euler showed that for a parabola, the time of flight is a function only of the sum of the lengths of the initial and final position vectors, $r_1 + r_2$, and the chord length, c.

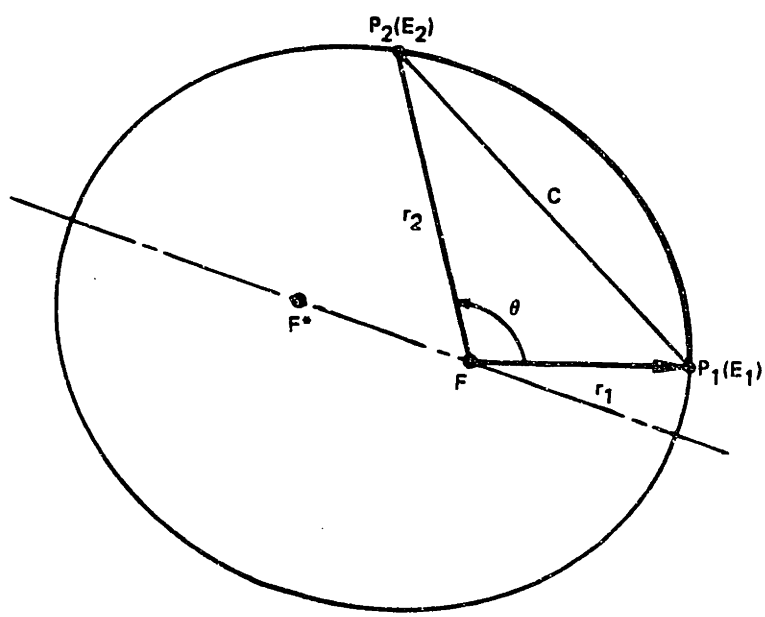


Figure 1. Geometry of Lambert's Problem

The next important contributor to this problem was the Swiss-German scholar Johann Heinrich Lambert. He independently derived Euler's solution for the parabolic orbit and extended it to include other conic sections such as the ellipse and the hyperbola. In his work *Insigniores orbitae Cometarum*

proprietates published in 1761, he presented what is now referred to as Lambert's theorem. He proved that the time of flight for an elliptic arc is a function only of the semi-major axis, a, and the quantities $r_1 + r_2$ and c. Lambert applied his work to the determination of the orbits of comets.[1] For his many contributions the two-point time-constrained orbital boundary value problem is commonly called Lambert's problem.

Lambert's proof of the theorem that bears his name was based on geometric considerations. A rigorous analytic treatment was presented by the French mathematician Joseph Louis Lagrange in 1778. He developed a simple expression for the time of flight as a function of the semi-major axis, a, and two parameters related to $r_1 + r_2$ and c. Both elliptic and hyperbolic orbits can be treated using these equations. Lagrange was also the first to note the geometric transformations of orbit implied by Lambert's theorem. One of these transformations is essential to the new solution presented in Chapter 3.

The single most important advance in solving Lambert's problem was presented by the German mathematician Carl Fredrich Gauss in his book *Theoria Motus* published in 1809. Gauss became interested in this problem in 1801 when he learned that astronomers had discovered a small body between the orbits of Jupiter and Mars.[2] This new body was believed to be the so-called "missing planet" predicted to lie between these two planets. Unfortunately, very little data were gathered about its orbit before it disappeared behind the sun. Gauss was intrigued by the problem of rediscovering this body when it emerged from behind the sun. Then only twenty-four years old, Gauss formulated the solution in terms of two unknown quantities simultaneously satisfying two algebraic equations at the solution point. The rediscovery of the asteroid Ceres, as it later came to be called, one year after its first observation at exactly the position predicted by Gauss' equations won him instant recognition as Europe's most outstanding mathematician.[3] Gauss' solution is significant not only for its analytic ingenuity but also for its computational ease and accuracy. A detailed derivation of Gauss' method is given in Chapter 2.

In the years following the publication of *Theoria Motus*, only minor improvements were suggested for the original method. Gauss' solution continued to be used as a tool for orbit determination in astronomy. With the advent of ballistic missiles and artificial satellites in the 1950's, the solution of Lambert's problem became the concern of engineers and scientists developing new aerospace technologies. Advanced guidance techniques called for fast and reliable solutions in cases not often considered by astronomers. Gauss' method was found to have several difficulties which precluded its implementation in guidance software. New solutions were developed, most of which involved

a Newton-Rapheson iteration process. (See, for example, Refs. 4–6.) The powered flight guidance program currently in use aboard the space shuttle employs such a scheme. Today, improved solutions of Lambert's problem are sought for application to spacecraft intercept and rendezvous problems, ballistic missile targeting problems, and interplanetary transfer problems.

Chapter 2

Gauss' Method

GAUSS' CLASSIC METHOD for solving Lambert's problem can be developed from the basic equations of two-body motion.[7] Although Gauss' method applies to both elliptic and hyperbolic orbits, the following derivation will consider only elliptic cases.

Kepler's equation relates the time of flight along an elliptic arc to the position as

$$\sqrt{\frac{\mu}{a^3}}(t-\tau) = E - e\sin E \tag{1}$$

where E is the eccentric anomaly, e is the orbital eccentricity, and τ is the time of passage through pericenter. Write Kepler's equation at the initial and final points as

$$\sqrt{\frac{\mu}{a^3}}(t_2-\tau) = E_2 - e\sin E_2 \qquad \text{and} \qquad \sqrt{\frac{\mu}{a^3}}(t_1-\tau) = E_1 - e\sin E_1.$$

Then, subtracting these two equations and using trigonometric identities gives

$$\begin{aligned}\sqrt{\frac{\mu}{a^3}}(t_2-t_1) &= E_2 - E_1 - e(\sin E_2 - \sin E_1)\\ &= 2\big(\tfrac{1}{2}(E_2-E_1) - e\cos\tfrac{1}{2}(E_2+E_1)\sin\tfrac{1}{2}(E_2-E_1)\big).\end{aligned}$$

Define the quantities ψ and ϕ as

$$\psi = \tfrac{1}{2}(E_2-E_1) \qquad \text{and} \qquad \cos\phi = e\cos\tfrac{1}{2}(E_2+E_1) \tag{2}$$

so that the time of flight equation can be written in terms of ψ and ϕ as

$$\sqrt{\frac{\mu}{a^3}}(t_2-t_1) = 2\psi - 2\sin\psi\cos\phi. \tag{3}$$

To eliminate $\cos\phi$ from the time of flight equation, the relations between the true anomaly, f, and eccentric anomaly

$$\sqrt{r}\cos\tfrac{1}{2}f = \sqrt{a(1-e)}\cos\tfrac{1}{2}E \qquad \text{and} \qquad \sqrt{r}\sin\tfrac{1}{2}f = \sqrt{a(1+e)}\sin\tfrac{1}{2}E \tag{4}$$

are used. The equations

$$\sqrt{r_1 r_2}\cos\tfrac{1}{2}f_1\cos\tfrac{1}{2}f_2 = a(1-e)\cos\tfrac{1}{2}E_2\cos\tfrac{1}{2}E_1$$

and

$$\sqrt{r_1 r_2}\sin\tfrac{1}{2}f_1\sin\tfrac{1}{2}f_2 = a(1+e)\sin\tfrac{1}{2}E_2\sin\tfrac{1}{2}E_1$$

are obtained by writing equations (4) at both points and multiplying. Adding the above relations and employing trigonmetric identities yields

$$\sqrt{r_1 r_2}\cos\tfrac{1}{2}\theta = a(\cos\psi - \cos\phi). \tag{5}$$

Substituting for $\cos\phi$ from equation (5) into the time of flight equation gives

$$\begin{aligned}\sqrt{\frac{\mu}{a^3}}(t_2 - t_1) &= 2\psi - 2\sin\psi\left(\cos\psi - \frac{\sqrt{r_1 r_2}}{a}\cos\tfrac{1}{2}\theta\right)\\ &= 2\psi - \sin 2\psi + 2\frac{\sqrt{r_1 r_2}}{a}\cos\tfrac{1}{2}\theta\sin\psi.\end{aligned}$$

The next step in the derivation is to eliminate the semi-major axis a from the time of flight equation. An expression for a is obtained by manipulating the equation of orbit for the ellipse

$$r = a(1 - e\cos E). \tag{6}$$

Writing equation (6) at the initial and final points and adding yields

$$\begin{aligned}r_1 + r_2 &= a(1 - e\cos E_1) + a(1 - e\cos E_2)\\ &= 2a(1 - \cos\psi\cos\phi).\end{aligned}$$

Rearranging to isolate a, gives

$$a = \frac{r_1 + r_2 - 2\sqrt{r_1 r_2}\cos\frac{1}{2}\theta\cos\psi}{2\sin^2\psi}. \tag{7}$$

At this point, Gauss introduced a quantity l defined by the equation

$$1 + 2l = \frac{r_1 + r_2}{2\sqrt{r_1 r_2}\cos\frac{1}{2}\theta}. \tag{8}$$

The expression for a now becomes

$$\begin{aligned}a &= \frac{\sqrt{r_1 r_2}\cos\frac{1}{2}\theta}{\sin^2\psi}\left((1 + 2l) - \cos\psi\right)\\ &= \frac{2\sqrt{r_1 r_2}\cos\frac{1}{2}\theta}{\sin^2\psi}(l + \sin^2\tfrac{1}{2}\psi).\end{aligned} \tag{9}$$

Using equation (9) to replace a in the time of flight equation gives

$$\frac{\sqrt{\mu}(t_2 - t_1)}{\left(2\sqrt{r_1 r_2}\cos\frac{1}{2}\theta\right)^{\frac{3}{2}}}\left(\frac{\sin^3\psi}{\left(l + \sin^2\frac{1}{2}\psi\right)^{\frac{3}{2}}}\right) = 2\psi - \sin 2\psi + \frac{\sin^3\psi}{l + \sin^2\frac{1}{2}\psi}$$

Rearranging,

$$\frac{\sqrt{\mu}(t_2 - t_1)}{\left(2\sqrt{r_1 r_2}\cos\frac{1}{2}\theta\right)^{\frac{3}{2}}} = \frac{2\psi - \sin 2\psi}{\sin^3\psi}\left(l + \sin^2\tfrac{1}{2}\psi\right)^{\frac{3}{2}} + \left(l + \sin^2\tfrac{1}{2}\psi\right)^{\frac{1}{2}}.$$

Again following Gauss, the quantities m* and y are defined as

$$m = \frac{\mu(t_2 - t_1)^2}{\left(2\sqrt{r_1 r_2}\cos\frac{1}{2}\theta\right)^3} \qquad \text{and} \qquad y^2 = \frac{m}{l + \sin^2\frac{1}{2}\psi}. \tag{10}$$

In terms of m and y, the time of flight equation takes the form

$$y = \left(\frac{2\psi - \sin 2\psi}{\sin^3\psi}\right)\frac{m}{y^2} + 1$$

or

$$y^3 - y^2 = m\left(\frac{2\psi - \sin 2\psi}{\sin^3\psi}\right). \tag{11}$$

Gauss proved that $(2\psi - \sin 2\psi)/\sin^3\psi$ is a hypergeometric function of

$$x = \sin^2\frac{1}{2}\psi = \sin^4\frac{1}{4}(E_2 - E_1)$$

and that it can be expressed as a continued fraction. Specifically,

$$\frac{2\psi - \sin 2\psi}{\sin^3\psi} = \tfrac{4}{3}F(3,1;\tfrac{5}{2};x) = \cfrac{1}{\frac{3}{4} - \cfrac{\frac{9}{10}x}{1 + \cfrac{\frac{2}{35}x}{1 - \cfrac{\frac{40}{63}x}{1 - \cfrac{\frac{4}{99}x}{1 - \cdots}}}}} \tag{12}$$

He next defined the function $\xi = \xi(x)$ by

$$x - \xi = \cfrac{x}{1 + \cfrac{\frac{2}{35}x}{1 - \cfrac{\frac{40}{63}x}{1 - \cfrac{\frac{4}{99}x}{1 - \cdots}}}}$$

from which

$$\xi = \cfrac{\frac{2}{35}x^2}{1 + \frac{2}{35}x - \cfrac{\frac{40}{63}x}{1 - \cfrac{\frac{4}{99}x}{1 - \cdots}}} \qquad \text{for} \qquad -\infty < x \leq 1 \tag{13}$$

Then the hypergeometric function can be written in terms of ξ as

$$\frac{2\psi - \sin 2\psi}{\sin^3\psi} = \frac{1}{\frac{3}{4} - \frac{9}{10}(x - \xi)}.$$

The time of flight equation becomes

$$y^2(y - 1) = \frac{m}{\frac{3}{4} - \frac{9}{10}(x - \xi)}$$

*Gauss originally denoted this quantity as m^2.

using the above expression for the hypergeometric function. But $x = m/y^2 - l$ from the definition of y, equation (10), so that

$$y^2(y-1) = \frac{m}{\frac{3}{4} - \frac{6}{10}(\frac{m}{y^2} - l - \xi)}.$$

Dividing both sides by y^2 yields

$$y - 1 = \frac{m}{\frac{9}{10}(\frac{5}{6} + l + \xi)y^2 - \frac{9}{10}m} = \frac{\frac{10}{9}h}{y^2 - h}$$

where h is defined by

$$h = \frac{m}{\frac{5}{6} + l + \xi}. \tag{14}$$

Finally, by clearing fractions, the time of flight equation becomes the cubic equation

$$y^3 - y^2 - hy - \tfrac{1}{9}h = 0. \tag{15}$$

The two equations which must be simultaneously satisfied at the solution point are

$$\frac{m}{y^2} = x + l \qquad \text{and} \qquad y^3 - y^2 - hy - \tfrac{1}{9}h = 0$$

where h depends on x through the function ξ. These equations are often referred to as the two equations of Gauss.

A similar derivation can be made for hyperbolic orbits beginning with the hyperbolic form of Kepler's equation which uses H, a quantity analogous to the eccentric anomaly. Now ψ and ϕ are defined as

$$\psi = \tfrac{1}{2}(H_2 - H_1) \qquad \text{and} \qquad \cosh\phi = e\cosh\tfrac{1}{2}(H_2 + H_1)$$

and x becomes

$$x = -\sinh^2\tfrac{1}{2}\psi = -\sinh^2\tfrac{1}{4}(H_2 - H_1).$$

The resulting equations are identical to those obtained for elliptic orbits. It is interesting to note that the parabolic orbit corresponds to $x = 0$. This follow directly from the equations of parabolic motion for which

$$m = l\left(1 + \tfrac{4}{3}l\right)^2 \tag{16}$$

In summary,

$$x = \begin{cases} -\sinh^2\tfrac{1}{4}(H_2 - H_1) & -\infty < x < 0 & \text{hyperbola} \\ 0 & & \text{parabola} \\ \sin^2\tfrac{1}{4}(E_2 - E_1) & 0 < x \leq 1 & \text{ellipse} \end{cases}$$

To simplify the notation, define two dimensionless parameters λ and T which depend on the given geometry and time of flight for the problem. Consider the triangle whose sides are r_1, r_2, and the chord c and let s be the semi-perimeter of the triangle

$$s = \tfrac{1}{2}(r_1 + r_2 + c). \tag{17}$$

Then λ is defined as

$$\lambda s = \sqrt{r_1 r_2}\cos\tfrac{1}{2}\theta \qquad \text{or} \qquad \lambda^2 = \frac{s-c}{s}. \tag{18}$$

a quantity which depends only on the given geometry and varies between $+1$ and -1 as the transfer angle θ varies from 0 to 2π. In particular, λ is zero for a transfer angle of $\theta = \pi$. It is positive for transfer angles less than π and negative for angles greater than π. Next, T is defined as

$$T = \sqrt{\frac{\mu}{(\frac{1}{2}s)^3}}(t_2 - t_1) = \sqrt{\frac{8\mu}{s^3}}(t_2 - t_1) \tag{19}$$

and depends both on the geometry (through s) and the specified flight time. These parameters can be taken as inputs for any particular problem since they are simple functions of r_1, r_2, and $(t_2 - t_1)$.

The quantities l and m defined by Gauss can be expressed as functions of λ and T. First, l is dependent on the given geometry and is a function of λ only. Specifically,

$$1 + 2l = \frac{r_1 + r_2}{2\sqrt{r_1 r_2}\cos\frac{1}{2}\theta} = \frac{2s - c}{2\lambda s} = \frac{1+\lambda^2}{2\lambda}$$

from which

$$l = \frac{(1-\lambda)^2}{4\lambda}. \tag{20}$$

Next, m clearly depends on the time of flight as well as the geometry and is a function of both λ and T.

$$m = \frac{\mu(t_2 - t_1)^2}{\left(2\sqrt{r_1 r_2}\cos\frac{1}{2}\theta\right)^3} = \frac{\frac{1}{8}s^3 T^2}{(2\lambda s)^3} = \frac{T^2}{64\lambda^3}. \tag{21}$$

The two equations developed by Gauss form the basis of a successive substitution procedure for solving Lambert's problem. An initial guess is chosen for x, such as $x = 0$. Then the continued fraction (13) is evaluated to obtain ξ. The coefficient h is found from ξ and the cubic equation (15) is solved for y. This value of y is used in the equation defining y written as $x = m/y^2 - l$ to obtain a new trial value of x. The entire process of evaluating ξ and h and solving for y to obtain a new x is repeated until x ceases to change within some specified tolerance. Figure 2 shows plots of T versus $-x$ for various values of λ.

The process of solving the cubic equation sometimes involves a choice among the real roots. Gauss presented guidelines for choosing the appropriate root which can be illustrated graphically.

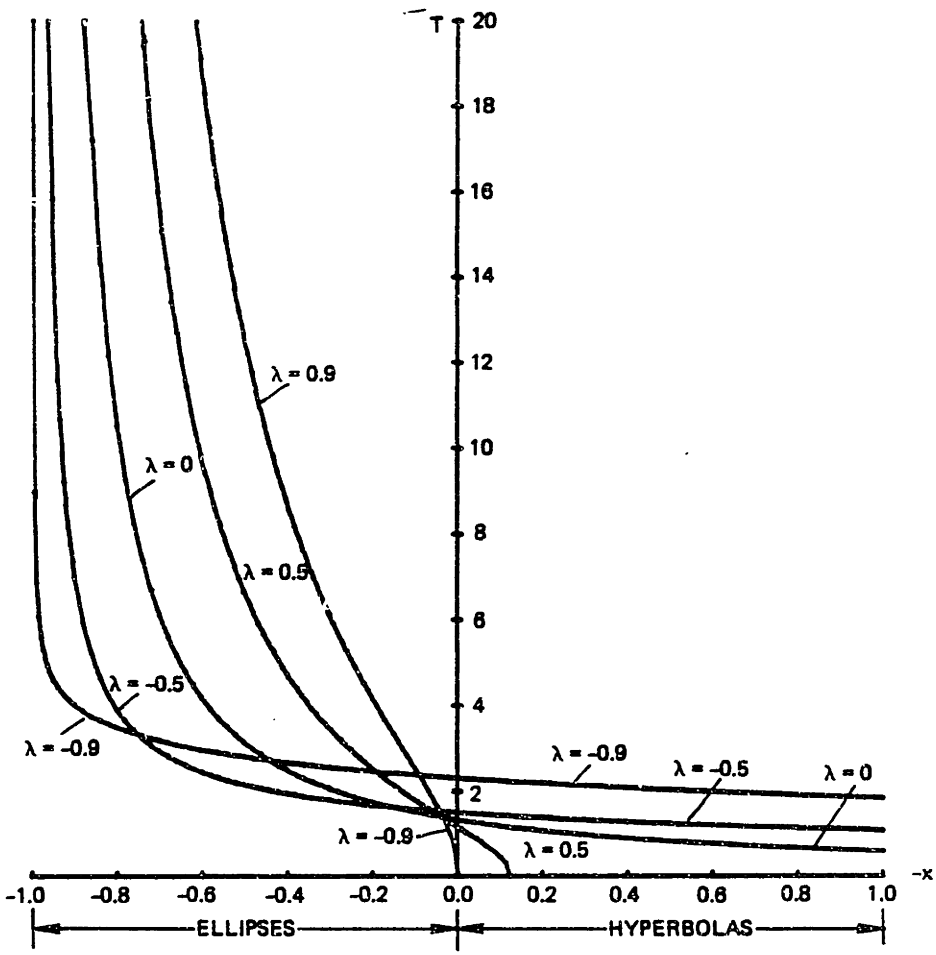


Figure 2. Time of Flight as a Function of $-x$ Using Gauss' Method

A plot of $y = y(h)$ obtained from the cubic written as

$$h = \frac{y^2(y-1)}{y+\frac{1}{9}}$$

is shown in Fig. 3. The branch of the cubic above the h axis where $y > 0$ represents cases of transfer angle $\theta < \pi$; the branches where $y < 0$ represent cases of transfer angle $\theta > \pi$.[8] For $h < \frac{1-\sqrt{5}}{6}$ and $0 < h < \frac{1+\sqrt{5}}{6}$, there is only one real root of the cubic equation. Gauss noted that for $\theta < \pi$, both h and y are positive quantities. There is no ambiguity in the choice of roots for these cases since there is only one positive real root for y. The cases of elliptic and hyperbolic orbits are separated by the point $y = 1 + \frac{4}{3}l$ which corresponds to the parabola. This point moves along the upper branch of the cubic from $y = 1$ as λ decreases from 1.

The situation is more complicated for $y < 0$. There is a singularity at the point $y = -\frac{1}{9}$. Also there are two negative real roots for y when $h > \frac{1+\sqrt{5}}{6}$. For elliptic orbits, Gauss showed that the desired root is the one for which $y < -\frac{1}{3}$. The point $y = -\frac{1}{3}$ is the limiting case of the parabola for $\lambda = 0$. Gauss limited his discussion of hyperbolic cases to small values of x for which he inferred that $y < \frac{1-\sqrt{5}}{6}$. He stated that at $y = \frac{1-\sqrt{5}}{6}$, $z = -.79858$ but

"we are far from wishing to extend our method to such great values of x."[9]

For hyperbolic cases where y approaches $\frac{1-\sqrt{5}}{6}$ the successive substitution process must be carefully applied. Some initial values of x may yield values of h which are negative or for which the cubic has no negative real root for y. It may not be possible to find a new trial value for x. Even if a new value for x is found, the iteration process may not return to the correct branch of the cubic equation.

The last step in the solution process is to determine the orbital elements from x and y. An expression for a is obtained from equation (9) using $\sin^2\psi = 4x(1-x)$

$$a = \frac{m\lambda s}{2x(1-x)y^2} = \frac{\lambda s(l+x)}{2x(1-x)}. \tag{22}$$

By combining the relations (4) at the initial and final points, $\sqrt{r_1 r_2}\sin\frac{1}{2}\theta = a\sqrt{1-e^2}\sin\psi$. An expression for p is found by squaring this and recalling that $p = a(1-e^2)$. Therefore,

$$p = \frac{r_1 r_2 \sin^2\frac{1}{2}\theta}{a\sin^2\psi}.$$

Substituting for a from equation (9),

$$p = \frac{\sqrt{r_1 r_2}\sin^2\frac{1}{2}\theta}{2\cos\frac{1}{2}\theta\left(l+\sin^2\frac{1}{2}\psi\right)}$$

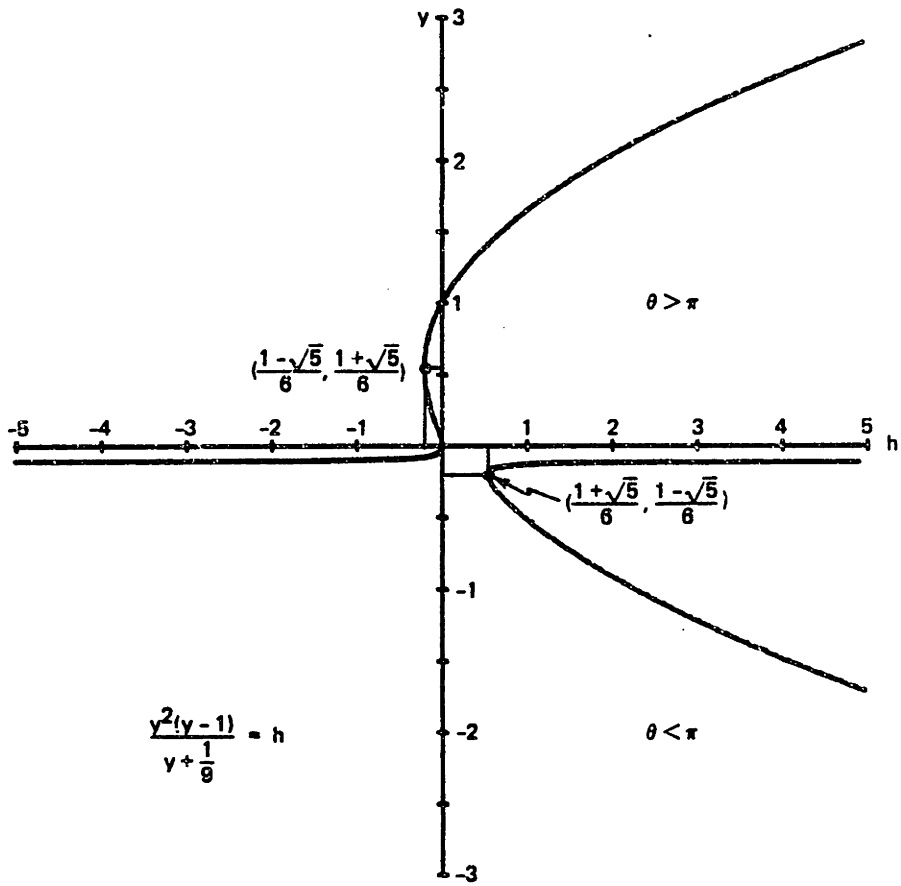


Figure 3. Cubic Equation of Gauss' Method

and using the definitions of λ, y, and m, the final expression becomes

$$p = \frac{y^2 r_1 r_2 \sin^2 \frac{1}{2}\theta}{2m\lambda s}. \tag{23}$$

This equation for p leads to an interesting physical interpretation of the variable y. Rearranging equation (23) to isolate y gives

$$y^2 = \frac{2mp\lambda s}{r_1 r_2 \sin^2 \frac{1}{2}\theta} = \frac{\mu p (t_2 - t_1)^2}{(r_1 r_2)^2 \sin^2 \theta}$$

or

$$y = \frac{\sqrt{\mu p}(t_2 - t_1)}{r_1 r_2 \sin \theta}. \tag{24}$$

Thus, from Kepler's second law, y is the ratio of the area of the sector bounded by $\mathbf{r}_1$, $\mathbf{r}_2$, and the elliptic arc to the area of the triangle formed by $\mathbf{r}_1$, $\mathbf{r}_2$, and the chord c.

A quantity of interest in most guidance problems is the velocity vector $\mathbf{v}_1$ of the two-body orbit corresponding to the initial position vector $\mathbf{r}_1$. Gauss did not give an equation for this quantity since he was involved with the determination of the orbits of comets and asteroids moving under the influence of gravity only. While the equations for a and p are sufficient for this application, the velocity $\mathbf{v}_1$ becomes important for controlling a spacecraft whose motion is influenced by its own propulsion system as well as gravity. It can be found from the given position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, and the orbital elements by introducing the Lagrangian coefficients. The position and velocity of any point on the orbit can be expressed in terms of the initial position and velocity by

$$\begin{bmatrix} \mathbf{r}_2 \\ \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} F & G \\ F_t & G_t \end{bmatrix} \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{v}_1 \end{bmatrix} \tag{25}$$

Where F, G, F_t, and G_t are called the Lagrangian coefficients. These coefficients are given in terms of the orbital parameter and the transfer angle θ below.[10]

$$\begin{aligned} F &= 1 - \frac{r_2}{p}(1 - \cos\theta) = 1 - \frac{2r_2}{p}\sin^2 \tfrac{1}{2}\theta \\ G &= \frac{r_1 r_2 \sin\theta}{\sqrt{\mu p}} \\ F_t &= \frac{\sqrt{\mu}}{r_1 p}\left(\frac{1}{\sqrt{\mu}}\mathbf{r}_1 \cdot \mathbf{v}_1 (1 - \cos\theta) - \sqrt{p}\sin\theta\right) \\ G_t &= 1 - \frac{r_1}{p}(1 - \cos\theta) \end{aligned} \tag{26}$$

Performing the matrix multiplication in equation (25) gives

$$\mathbf{r}_2 = F\mathbf{r}_1 + G\mathbf{v}_1$$

or, isolating $\mathbf{v}_1$,

$$\mathbf{v}_1 = \frac{\mathbf{r}_2 - F\mathbf{r}_1}{G} \tag{27}$$

The desired equation for $\mathbf{v}_1$ is obtained by writing the Lagrangian coefficients F and G in terms of the input quantities and the final value of y from Gauss' method. First substitute from equation (23) for p into the relation for F.

$$F = 1 - \frac{2r_2 \sin^2 \frac{1}{2}\theta}{p} = 1 - \frac{4m\lambda s}{y^2 r_1}$$

Next use equation (24) for y to find

$$G = \frac{(t_2 - t_1)}{y}.$$

Substituting these two relations for F and G into equation (27) for $\mathbf{v}_1$ gives

$$\begin{aligned} \mathbf{v}_1 &= \frac{y}{(t_2 - t_1)}\left(\mathbf{r}_2 - \left(1 - \frac{4m\lambda s}{y^2 r_1}\right)\mathbf{r}_1\right) \\ &= \frac{4m\lambda s - y^2 r_1}{y r_1 (t_2 - t_1)}\mathbf{r}_1 + \frac{y}{(t_2 - t_1)}\mathbf{r}_2 \end{aligned} \tag{28}$$

Chapter 3

NEW METHOD

THE NEW METHOD of solving Lambert's problem presented here is based upon the geometric transformations of orbit implied by Lambert's theorem. This theorem can be expressed as

$$\sqrt{\mu}(t_2 - t_1) = f(a, r_1 + r_2, c) \tag{29}$$

and implies that the shape of the orbit can be changed without changing the time of flight provided a, $r_1 + r_2$, and c are fixed.

3.1 Transformation of Orbit

If the termini of the position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, denoted P_1 and P_2, are fixed, the chord length c is also fixed. The occupied and vacant foci F and F^* can be moved to change the shape of the orbit. The conditions that $r_1 + r_2$ and a be held constant constrain the motion of F and F^* to lie along ellipses with foci at P_1 and P_2 whose major axes are $r_1 + r_2$ and $4a - (r_1 + r_2)$, respectively. Figure 4 shows a transformation of elliptic orbits from foci F and F^* to F_i and F_i^* as well as the ellipses representing allowable positions of F and F^*.

Before describing the particular transformation of interest in the current problem, it is necessary to define the mean point of the orbital boundary value problem. The mean point, corresponding to radius r_o, is that point along the arc from P_1 to P_2 where the velocity vector, or orbital tangent, is parallel to the chord. It is easily shown that the mean points of all possible orbits connecting P_1 and P_2 are collinear. Furthermore, the eccentric anomaly of the mean point for an elliptic orbit is the arithmetic mean of the eccentric anomalies of the initial and final points. A similar relation holds for hyperbolic orbits, so that

$$E_o = \tfrac{1}{2}(E_2 + E_1) \qquad \text{and} \qquad H_o = \tfrac{1}{2}(H_2 + H_1).$$

Using the equations of orbit, the mean point radius is written

$$r_o = a\big(1 - e\cos\tfrac{1}{2}(E_2 + E_1)\big) \qquad \text{and} \qquad r_o = a\big(1 - e\cosh\tfrac{1}{2}(H_2 + H_1)\big). \tag{30}$$

The mean point of the parabolic orbit, r_{op}, connecting P_1 and P_2 can be simply expressed as

$$r_{op} = \tfrac{1}{4}\big(r_1 + r_2 + 2\sqrt{r_1 r_2}\cos\tfrac{1}{2}\theta\big) \tag{31}$$

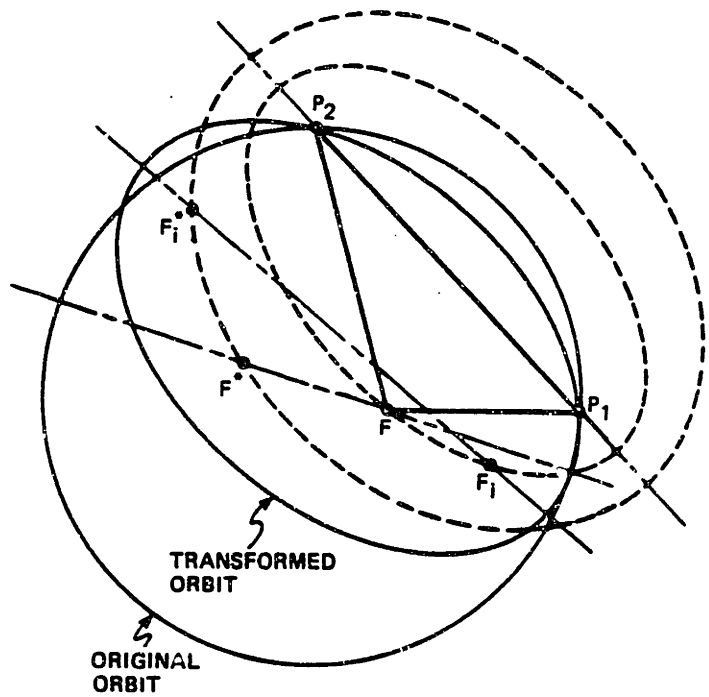


Figure 4. Transformations of Lambert's Problem

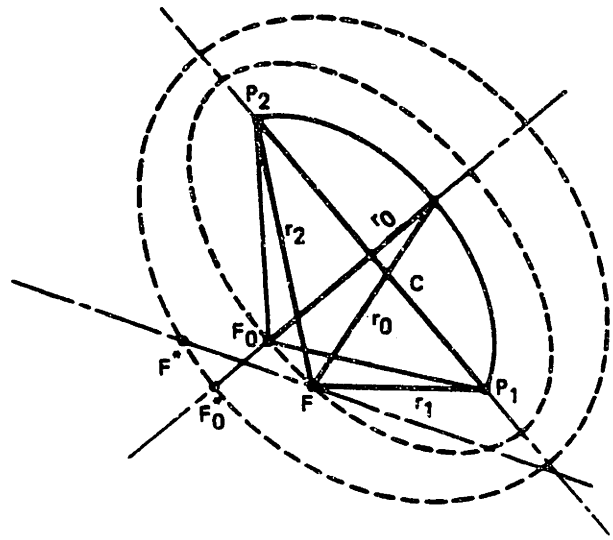
so that the mean point radius of any orbit can be written in terms of r_{op} as follows[11]

$$r_o = \begin{cases} r_{op} \sec^2 \frac{1}{2}\psi & \text{ellipse} \\ r_{op} \operatorname{sech}^2 \frac{1}{2}\psi & \text{hyperbola} \end{cases} \tag{32}$$

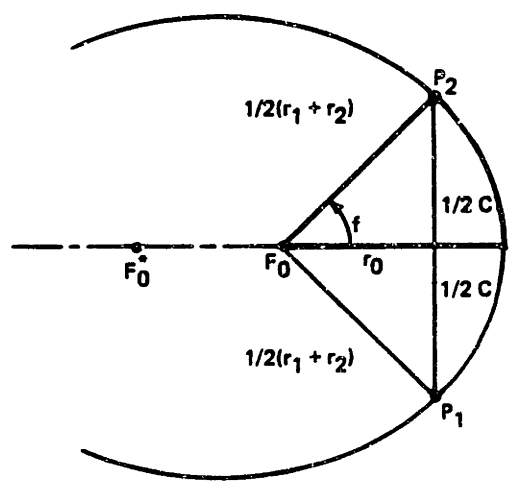
The transformation of orbit is now chosen so that the mean point radius r_o is perpendicular to the chord c as shown in Fig. 5. For an ellipse the mean point of the transformed orbit will be an apse, say pericenter. The transformed initial and final position vectors are therefore of equal length $\frac{1}{2}(r_1 + r_2)$. Also, the new transfer angle is twice the true anomaly f of point P_2. Since the time of flight is constant and the mean point is pericenter, the time to travel along the transformed arc from r_o to P_2 is one half of the original flight time between P_1 and P_2. The simple form of Kepler's equation can be written for the arc from r_o to P_2

$$\frac{1}{2}\sqrt{\frac{\mu}{a^3}}(t_2 - t_1) = E - e_o \sin E \tag{33}$$

where e_o is the eccentricity and E the eccentric anomaly at P_2 for the transformed orbit.[12]



5(a)



5(b)

Figure 5. Transformed Version of Lambert's Problem

Several important quantities are invariant under the orbital transformation which permit the quantities describing the transformed problem to be related to those describing the original problem. One obvious invariant is the semi-perimeter s which depends only on $r_1 + r_2$ and c. Similarly, λ and T are unchanged since λ can be expressed in terms of s and c and T in terms of s and the time of flight. Lagrange showed that the expressions ψ and $\cos\phi$ can be expressed in terms of s

and a. Hence, ψ and $\cos\phi$ are invariants. Since r_o is pericenter, the eccentric anomaly at P_2 for the transformed orbit is one half the difference between the eccentric anomalies at P_1 and P_2. Also, the eccentric anomaly E_o of the mean point for the transformed orbit is zero. Thus,

$$\psi = \tfrac{1}{2}(E_2 - E_1) = E \qquad \text{and} \qquad \cos\phi = e\cos\tfrac{1}{2}(E_2 + E_1) = e_o. \tag{34}$$

Since $\sqrt{r_1 r_2}\cos\frac{1}{2}\theta = \pm\sqrt{s(s-c)}$, another invariant is

$$\sqrt{r_1 r_2}\cos\tfrac{1}{2}\theta = \tfrac{1}{2}(r_1 + r_2)\cos f \tag{35}$$

Two other invariants are the mean point radius and the mean point radius of the parabolic orbit. Specifically,

$$r_o = a\big(1 + e\cos\tfrac{1}{2}(E_2 + E_1)\big) = a(1 + e_o) \tag{36}$$

and

$$r_{op} = \tfrac{1}{4}\big(r_1 + r_2 + 2\sqrt{r_1 r_2}\cos\tfrac{1}{2}\theta\big) = \tfrac{1}{2}(r_1 + r_2)\cos^2\tfrac{1}{2}f \tag{37}$$

Two important quantities which are *not* invariant under transformation are the orbital eccentricity and the parameter. Therefore, $e_o \neq e$ and $p_o \neq p$.

3.2 Two Equations of the Transformed Problem

Two equations analogous to the two equations of Gauss can be derived from the simple form of Kepler's equation for the transformed problem and the invariants of transformation. Rewriting equation (33) and substituting r_o/a for $(1 - e_o)$ gives

$$\frac{1}{2}\sqrt{\frac{\mu}{a^3}}(t_2 - t_1) = \psi - \sin\psi + \frac{r_o}{a}\sin\psi.$$

Expressing r_o in terms of r_{op} and using trigonometric identities, there results

$$\frac{1}{2}\sqrt{\frac{\mu}{a^3}}(t_2 - t_1) = \psi - \sin\psi + \frac{2r_{op}}{a}\tan\tfrac{1}{2}\psi. \tag{38}$$

The semi-major axis a can be written as

$$a = \frac{r_o}{1 - e_o} = \frac{r_{op}(1 + \tan^2\frac{1}{2}\psi)}{1 - e_o}.$$

To eliminate e_o from this expression, use the equation of orbit for the transformed problem $r(1 + e_o\cos f) = p_o$ to obtain

$$\tfrac{1}{2}(r_1 + r_2)(1 + e_o\cos f) = r_o(1 + e_o).$$

Simplifying,

$$e_o = \frac{\frac{1}{2}(r_1 + r_2) - r_o}{r_o - \frac{1}{2}(r_1 + r_2)\cos f} \tag{39}$$

so that

$$1 - e_o = \frac{2r_o - \frac{1}{2}(r_1 + r_2)(1 + \cos f)}{r_o - \frac{1}{2}(r_1 + r_2)\cos f}.$$

Substituting from equations (32) and (37) for r_o and $\frac{1}{2}(r_1 + r_2)$ yields

$$1 - e_o = \frac{2\tan^2 \frac{1}{2}\psi}{\tan^2 \frac{1}{2}\psi + \tan^2 \frac{1}{2}f}. \tag{40}$$

Using this relation in the equation for a, there results

$$a = \frac{r_{op}(1 + \tan^2 \frac{1}{2}\psi)(\tan^2 \frac{1}{2}f + \tan^2 \frac{1}{2}\psi)}{2\tan^2 \frac{1}{2}\psi}.$$

Next, the quantity

$$l = \tan^2 \tfrac{1}{2}f \tag{41}$$

is introduced so that

$$a = \frac{r_{op}(1 + \tan^2 \frac{1}{2}\psi)(l + \tan^2 \frac{1}{2}\psi)}{2\tan^2 \frac{1}{2}\psi}. \tag{42}$$

Substituting for a from equation (42) in the time of flight equation,

$$\frac{1}{2}\sqrt{\mu}(t_2 - t_1)\left(\frac{2\tan^2 \frac{1}{2}\psi}{r_{op}(1 + \tan^2 \frac{1}{2}\psi)(l + \tan^2 \frac{1}{2}\psi)}\right)^{\frac{3}{2}} = \psi - \sin\psi + \frac{4\tan^3 \frac{1}{2}\psi}{(1 + \tan^2 \frac{1}{2}\psi)(l + \tan^2 \psi)}.$$

Rearranging,

$$\sqrt{\frac{2\mu}{{r_{op}}^3}}(t_2 - t_1) = \left(\frac{\psi - \sin\psi}{\tan^3 \frac{1}{2}\psi}\right)(1 + \tan^2 \tfrac{1}{2}\psi)^{\frac{3}{2}}(l + \tan^2 \tfrac{1}{2}\psi)^{\frac{3}{2}} + 4(1 + \tan^2 \tfrac{1}{2}\psi)^{\frac{1}{2}}(l + \tan^2 \tfrac{1}{2}\psi)^{\frac{1}{2}}.$$

The quantities m and y are now defined by

$$m = \frac{\mu(t_2 - t_1)^2}{8{r_{op}}^3} \qquad \text{and} \qquad y^2 = \frac{m}{(1 + \tan^2 \frac{1}{2}\psi)(l + \tan^2 \frac{1}{2}\psi)}. \tag{43}$$

Finally, using these definitions, the time of flight equation becomes

$$y^3 - y^2 = \frac{1}{4}m\left(\frac{\psi - \sin\psi}{\tan^3 \frac{1}{2}\psi}\right).$$

At this point, the quantity x is defined as $x = \tan^2 \frac{1}{2}\psi$. In terms of x, the two equations which are analogous to the two equation of Gauss are

$$y^2 = \frac{m}{(1 + x)(l + x)} \tag{44}$$

and

$$y^3 - y^2 = \frac{1}{4}m\left(\frac{\psi - \sin\psi}{\tan^3 \frac{1}{2}\psi}\right). \tag{45}$$

The quantity $(\psi - \sin\psi)/\tan^3 \frac{1}{2}\psi$ can be shown to be a hypergeometric function of x.

$$\frac{\psi - \sin\psi}{\tan^3 \frac{1}{2}\psi} = \frac{4}{3} F(\tfrac{3}{2}, 2; \tfrac{5}{2}; -x)$$

or, from the properties of hypergeometric functions[13]

$$\frac{\psi - \sin\psi}{\tan^3 \frac{1}{2}\psi} = -4 \frac{d}{dx} F(\tfrac{1}{2}, 1; \tfrac{3}{2}; -x) \tag{46}$$

where

$$F(\tfrac{1}{2}, 1; \tfrac{3}{2}; -x) = \frac{\tan^{-1}\sqrt{x}}{\sqrt{x}} = \cfrac{1}{1 + \cfrac{x}{3 + \cfrac{4x}{5 + \cfrac{9x}{7 + \cfrac{16x}{9 + \cdots}}}}} \tag{47}$$

3.3 Introduction of a Free Parameter

The new method now takes a different path from that originally followed by Gauss. A great advantage of the simple form of Kepler's equation, first noted by Gauss, is that a free parameter can be introduced. While Gauss used this property in his solution of Kepler's equation only, it can be exploited for Lambert's problem under the present orbital transformation. With the introduction of the free parameter β, Kepler's equation becomes

$$\frac{1}{2}\sqrt{\frac{\mu}{a^3}}(t_2 - t_1) = \big(1 + \beta(1 - e_o)\big)P + (1 - e_o)Q \tag{48}$$

where

$$P = \psi - \sin\psi \qquad \text{and} \qquad Q = \sin\psi - \beta P. \tag{49}$$

From equation (46) and basic trigonometric identities, P and Q are written as

$$P = -4\tan^3 \tfrac{1}{2}\psi \frac{dF}{dx} \qquad \text{and} \qquad Q = 2\tan\tfrac{1}{2}\psi\left(\frac{1}{1+x} + 2\beta x \frac{dF}{dx}\right)$$

where $F = F(\frac{1}{2}, 1; \frac{3}{2}; -x)$. But $1 - e_o = 2x/(l + x)$ from equation (40), so that

$$(1 - e_o)Q = \frac{4\tan^3 \frac{1}{2}\psi}{(1+x)(l+x)}\left(1 + 2\beta x(1+x)\frac{dF}{dx}\right).$$

The parameter h_1 is defined as

$$h_1 = 2\beta x(1+x)\frac{dF}{dx}. \tag{50}$$

Hence,

$$(1-e_o)Q = \frac{4\tan^3\frac{1}{2}\psi}{(1+x)(l+x)}(1+h_1) = \frac{4y^2}{m}\tan^3\tfrac{1}{2}\psi(1+h_1) \tag{51}$$

and

$$\big(1+\beta(1-e_o)\big)P = -4\tan^3\tfrac{1}{2}\psi\left(\frac{dF}{dx}+\frac{h_1}{(1+x)(l+x)}\right) \tag{52}$$

Substituting these two terms into Kepler's equation gives

$$\frac{1}{2}\sqrt{\frac{\mu}{a^3}}(t_2-t_1) = 4\tan^3\tfrac{1}{2}\psi\left(\frac{y^2}{m}(1+h_1)-\frac{dF}{dx}-\frac{h_1}{(1+x)(l+x)}\right).$$

Using $a = \big(r_{op}(1+x)(l+x)\big)/2x$, equation (42), and the definitions of y and m,

$$\frac{1}{2}\sqrt{\frac{\mu}{a^3}}(t_2-t_1) = \frac{\sqrt{\mu}(t_2-t_1)}{2}\left(\frac{2x}{r_{op}(1+x)(l+x)}\right)^{\frac{3}{2}} = \frac{4y^3}{m}\tan^3\tfrac{1}{2}\psi.$$

With this final substitution, Kepler's equation becomes the cubic

$$\frac{y^3}{m} = \frac{y^2}{m}(1+h_1)-\frac{dF}{dx}-\frac{h_1}{(1+x)(l+x)} \qquad \text{or} \qquad y^3-(1+h_1)y^2+m\left(\frac{dF}{dx}+\frac{h_1}{(1+x)(l+x)}\right) = 0. \tag{53}$$

Note that if the free parameter $\beta = 0$, $h_1 = 0$ and equation (53) reduces to the previous cubic equation (45). To further simplify the notation, define

$$h_2 = -m\left(\frac{dF}{dx}+\frac{h_1}{(1+x)(l+x)}\right) \tag{54}$$

so that the cubic equation is written as

$$y^3-(1+h_1)y^2-h_2 = 0. \tag{55}$$

3.4 Choosing the Free Parameter

The two equations analogous to those of Gauss expressed in terms of the free parameter are

$$y^2 = \frac{m}{(1+x)(l+x)} \qquad \text{and} \qquad y^3-(1+h_1)y^2-h_2 = 0$$

where the definitions of h_1 and h_2 contain the free parameter β. It is desired to choose β so that a successive substitution iteration will converge as rapidly as possible. To this end, β is chosen so that the first derivative, dy/dx, of the cubic equation is zero at the solution point. (A detailed motivation for this choice will be presented in a later section.) Differentiating equation (53) yields

$$y\big(3y-2(1+h_1)\big)\frac{dy}{dx}-y^2\frac{dh_1}{dx}+m\frac{d^2F}{dx^2}+\frac{m}{(1+x)(l+x)}\frac{dh_1}{dx}+mh_1\frac{d}{dx}\left(\frac{1}{(1+x)(l+x)}\right) = 0.$$

The second and fourth terms are eliminated since equation (44) holds at the solution point. Setting $dy/dx = 0$ gives

$$\frac{d^2F}{dx^2} + h_1 \frac{d}{dx}\left(\frac{1}{(1+x)(l+x)}\right) = 0$$

as the equation determining h_1 and hence β. Performing the differentiation in the second term,

$$\frac{d^2F}{dx^2} - \frac{h_1(1+2x+l)}{(1+x)^2(l+x)^2} = 0$$

or

$$(1+x)^2\frac{d^2F}{dx^2} - \frac{h_1(1+2x+l)}{(l+x)^2} = 0 \tag{56}$$

To find the derivatives of F, first note that F satisfies Gauss' differential equation for hypergeometric functions:

$$2x(1+x)\frac{d^2F}{dx^2} + (3+5x)\frac{dF}{dx} + F = 0. \tag{57}$$

Replacing the second derivative of F in equation (56) using equation (57) gives

$$(1+x)\left((3+5x)\frac{dF}{dx} + F\right) + \frac{2x(1+2x+l)}{(l+x)^2}h_1 = 0. \tag{58}$$

To evaluate the first derivative of F, two other properties of hypergeometric functions are employed.[14] First,

$$\frac{dF}{dx} = \frac{1}{2x}\left(F(1,\tfrac{3}{2};\tfrac{3}{2};-x) - F\right).$$

Next,

$$F(1,\gamma;\gamma;-x) = \frac{1}{1+x} \qquad \text{for any } \gamma$$

so that the first derivative of F is

$$\frac{dF}{dx} = \frac{1}{2x}\left(\frac{1}{1+x} - F\right). \tag{59}$$

Define the functions G and ξ using the continued fraction for F, equation (47),

$$F = \frac{1}{1+xG} \qquad \text{and} \qquad G = \frac{1}{3+4x\xi}. \tag{60}$$

Hence,

$$G = \cfrac{1}{3+\cfrac{4x}{5+\cfrac{9x}{7+\cfrac{16x}{9+\cdots}}}}$$

and

$$\xi = \cfrac{1}{5+\cfrac{9x}{7+\cfrac{16x}{9+\cfrac{25x}{11+\cdots}}}} = \cfrac{\frac{1}{5}}{1+\cfrac{\frac{9}{35}x}{1+\cfrac{\frac{16}{63}x}{1+\cfrac{\frac{25}{99}x}{1+\cdots}}}} \qquad \text{for} \qquad -1 \le x < \infty \tag{61}$$

Substituting for F in terms of G in equation (59) yields

$$\frac{dF}{dx} = \frac{1}{2x}\left(\frac{1}{1+x} - \frac{1}{1+xG}\right) = -\frac{1}{2}\frac{(1-G)F}{1+x}.$$

But

$$1 - G = 1 - \frac{1}{3+4\xi} = 2(1+2x\xi)G$$

and

$$FG = \left(\frac{1}{1+xG}\right)\left(\frac{1}{3+4x\xi}\right) = \frac{1}{\left(1+\frac{x}{3+4x\xi}\right)(3+4x\xi)} = \frac{1}{3+x(1+4\xi)},$$

so that

$$\frac{dF}{dx} = \frac{-(1+2x\xi)FG}{1+x} = \frac{-(1+2x\xi)}{(1+x)(3+x(1+4\xi))}. \tag{62}$$

The first term of equation (58) can be written as

$$(1+x)\left((3+5x)\frac{dF}{dx} + F\right) = (1+x)\left(F - \frac{(3+5x)(1+2x\xi)FG}{1+x}\right)$$

using equation (62) for dF/dx. Simplifying,

$$\begin{aligned}(1+x)\left((3+5x)\frac{dF}{dx} + F\right) &= FG\big((1+x)(3+4x\xi) - (1+2x\xi)(3+5x)\big)\\ &= -2xFG(1+\xi+3x\xi)\\ &= -\frac{2x(1+\xi(1+3x))}{(3+x(1+4\xi))}\end{aligned}$$

Now the determining equation for h_1 becomes

$$-\frac{2x(1+\xi(1+3x))}{(3+x(1+4\xi))} + \frac{2x(1+2x+l)h_1}{(l+x)^2} = 0.$$

Hence

$$h_1 = \frac{(l+x)^2(1+\xi(1+3x))}{(1+2x+l)(3+x(1+4\xi))}. \tag{63}$$

This equation implicitly contains a choice of the free parameter β. An explicit expression for β can be obtained by substituting for dF/dx from equation (62) in equation (50) for h_1 and setting the result equal to that of equation (63) above. This step is not necessary since the quantity of interest here is the cubic coefficient h_1. Finally, h_2 is found from equation (62) for dF/dx and the above equation for h_1.

$$h_2 = \frac{m}{1+x}\left(\frac{1+2x\xi}{3+x(1+4\xi)} - \frac{(l+x)(1+\xi(1+3x))}{(1+2x+l)(3+x(1+4\xi))}\right)$$

or, simply

$$h_2 = \frac{m(1+(x-l)\xi)}{(1+2x+l)(3+x(1+4\xi))}. \tag{64}$$

The derivation can be extended to include parabolic and hyperbolic orbits by using

$$z = \begin{cases} -\tanh^2 \frac{1}{4}(H_2 - H_1) & -1 \le x < 0 & \text{hyperbola} \\ 0 & & \text{parabola} \\ \tan^2 \frac{1}{4}(E_2 - E_1) & 0 < x < \infty & \text{ellipse} \end{cases}$$

Note that for parabolic motion,

$$m = l(1 + \tfrac{1}{3}l)^2. \tag{65}$$

The new l and m are similar to those defined by Gauss in that they are functions of the given geometry and flight time for a particular problem. To express l and m in terms of λ and T, recall that $\sqrt{r_1 r_2}\cos\frac{1}{2}\theta$ is a transformation invariant. Isolate $\cos f$ using equation (35) so that

$$\cos f = \frac{\sqrt{r_1 r_2}\cos\frac{1}{2}\theta}{\frac{1}{2}(r_1 + r_2)} = \frac{2\lambda s}{2s - c}.$$

Using trigonometric identities to relate $\cos f$ to $\tan^2 \frac{1}{2}f$ gives

$$\cos f = \frac{1 - \tan^2 \frac{1}{2}f}{1 + \tan^2 \frac{1}{2}f} = \frac{1 - l}{1 + l} = \frac{2\lambda s}{2s - c}.$$

Rearranging to isolate l,

$$l = \frac{s - c + s(1 - 2\lambda)}{s - c + s(1 + 2\lambda)} = \frac{\lambda^2 + 1 - 2\lambda}{\lambda^2 + 1 + 2\lambda} = \left(\frac{1 - \lambda}{1 + \lambda}\right)^2. \tag{66}$$

To express m is terms of λ and T, rewrite equation (37) for r_{op} in terms of λ and s.

$$r_{op} = \tfrac{1}{4}(2s - c + 2\lambda s) = \tfrac{1}{4}s(1 + \lambda)^2.$$

Using the above equation for r_{op} in the defining equation of m yields

$$m = \frac{\mu(t_2 - t_2)^2}{8(\frac{1}{64})s^3(1 + \lambda)^6} = \frac{8\mu(t_2 - t_1)^2}{s^3(1 + \lambda)^6} = \frac{T^2}{(1 + \lambda)^6}. \tag{67}$$

3.5 The New Iteration Process

A new iteration process can be formulated based on equations (44) and (55) . First an initial value is chosen for x. Then the continued fraction, equation (61), is used to obtain ξ and the two coefficients of the cubic, h_1 and h_2, are computed from equations (63) and (64) using ξ. Next, the cubic equation (55) is solved for y and a new value of x is obtained from equation (44) . While Gauss' defining equation for y contained a linear term in x, equation (44) is quadratic in the new x. There are two

values of x corresponding to a value of y^2. Specifically,

$$x = \frac{1}{2}\left(-(1+l) \pm \sqrt{(1-l)^2 + \frac{4m}{y^2}}\right).$$

The choice of sign in the above relation becomes apparent when it is recalled that the range of x is $-1 \leq x < \infty$. Therefore, the + sign is chosen and the new value of x computed from

$$x = \frac{1}{2}\left(\sqrt{(1-l)^2 + \frac{4m}{y^2}} - (1+l)\right)$$

The iteration continues as outlined above until the values of x and y converge within some specified tolerance. Figure 6 shows plots of T versus x for the new method.

It now remains to find expressions for the orbital elements in terms of the new m, x, and y. A simple relation for the semi-major axis a is obtained from equation (42), $a = r_{op}(1+x)(l+x)/2x$. Recalling that $(1+x)(l+x) = m/y^2$ and $r_{op} = \frac{1}{4}s(1+\lambda)^2$, this becomes

$$a = \frac{r_{op}m}{2xy^2} = \frac{ms(1+\lambda)^2}{8xy^2}. \tag{68}$$

The expression for the parameter p is developed using two results from the derivation of Gauss' method, namely $\sqrt{r_1 r_2}\sin\frac{1}{2}\theta = a\sqrt{1-e^2}\sin\psi$ from which

$$p = \frac{r_1 r_2 \sin^2\frac{1}{2}\theta}{a\sin^2\psi}.$$

Using basic trigonmetric identities, replace $\sin^2\psi$ with $4x/(1+x)^2$, where $x = \tan^2\frac{1}{2}\psi$, to obtain

$$p = \frac{r_1 r_2 (1+x)^2 \sin^2\frac{1}{2}\theta}{4ax}.$$

Substituting from the previously derived equation for a gives

$$p = \frac{2r_1 r_2 y^2 (1+x)^2 \sin^2\frac{1}{2}\theta}{ms(1+\lambda)^2}. \tag{69}$$

It is important to note that equations (68) and (69) are expressions for the semi-major axis and the parameter of the *original* orbit. The distinction need not be made for a since it is a transformation invariant. However, the parameters of the original orbit and the transformed orbit are *not* the same.

An expression for the initial velocity vector $\mathbf{v}_1$ can be found using the equation for p. From the previous chapter, the initial velocity vector can be written as

$$\mathbf{v}_1 = \frac{\mathbf{r}_2 - F\mathbf{r}_1}{G} \tag{27}$$

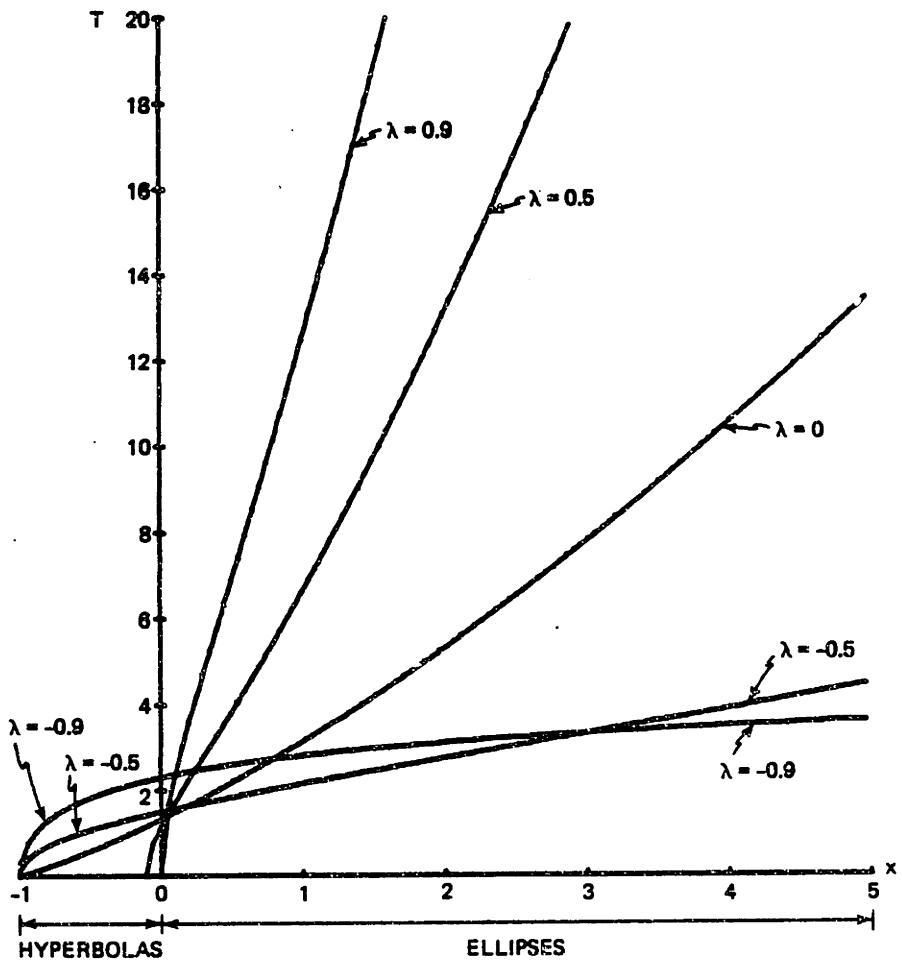


Figure 6. Time of Flight as a Function of x Using the New Method

where

$$F = 1 - \frac{2r_2}{p}\sin^2 \tfrac{1}{2}\theta \qquad \text{and} \qquad G = \frac{r_1 r_2 \sin\theta}{\sqrt{\mu p}}.$$

Substituting from equation (69) for p in the equations for F and G gives

$$G = \sqrt{\frac{2ms^3}{\mu}}\,\frac{(1+\lambda)}{y(1+z)} \qquad \text{and} \qquad F = 1 - \frac{ms(1+\lambda)^2}{r_1 y^2(1+z)^2}.$$

The equation for G is simplified by substituting from equation (67) for m

$$G = \frac{4(t_2 - t_1)}{y(1+z)(1+\lambda)^2}.$$

Putting the new relations for F and G into equation (27) for $\mathbf{v}_1$ yields

$$\begin{aligned}\mathbf{v}_1 &= \frac{y(1+z)(1+\lambda)^2}{4(t_2-t_1)}\left(\mathbf{r}_2 - \left(1 - \frac{ms(1+\lambda)^2}{r_1 y^2(1+z)^2}\right)\mathbf{r}_1\right)\\ &= \frac{(1+\lambda)^2}{4(t_2-t_1)}\left(y(1+z)(\mathbf{r}_2 - \mathbf{r}_1) + \frac{ms(1+\lambda)^2}{r_1 y(1+z)}\mathbf{r}_1\right)\end{aligned} \tag{70}$$

Chapter 4

Details of the New Iteration

THE TWO PARTS of the iteration procedure using either Gauss' method or the new method which require the greatest amount of computation are the evaluation of the continued fraction for ξ and the solution of the cubic equation. Various techniques for accomplishing these tasks in the new method are presented in this chapter. First, a brief description of a "top-down" method for evaluating continued fractions is given. Next, a method of improving the convergence of the continued fraction for large values of x is developed. Finally, three methods for solving the cubic equation are discussed.

4.1 Evaluating the Continued Fraction

The method of evaluating the continued fraction presented below is based on a formula attributed to Euler which converts the fraction into an equivalent series.[15] Consider the continued fraction

$$c = \cfrac{a_1}{b_1 - \cfrac{a_2}{b_2 - \cfrac{a_3}{b_3 - \cfrac{a_4}{b_4 - \cdots}}}}. \tag{71}$$

A recursive algorithm for evaluating c to any desired precision is described by the equations

$$\begin{aligned} \delta_n &= \frac{1}{1 - \frac{a_n}{b_{n-1} b_n} \delta_{n-1}} \\ u_n &= u_{n-1}(\delta_n - 1) \qquad\qquad \text{for } n = 2, 3, 4, \ldots \\ \frac{p_n}{q_n} &= \sum_{i=1}^{n} u_i \end{aligned} \tag{72}$$

where $\delta_1 = 1$ and $u_1 = a_1/b_1$. The iteration is continued until that value of n is reached for which u_n is within some specified tolerance at which point $c \approx p_n/q_n$. This is called a "top-down" method since each step n of the iteration reaches down one level further in the continued fraction for the values of a_n and b_n.

The iteration can be further simplified if a formula for generating a_n and b_n from the previous values a_{n-1} and b_{n-1} can be found. Consider the expression for $\xi(x)$ repeated below

$$\xi = \cfrac{\frac{1}{5}}{1 + \cfrac{\frac{9}{35}x}{1 + \cfrac{\frac{16}{63}x}{1 + \cfrac{\frac{25}{99}x}{1 + \cdots}}}}. \tag{61}$$

Clearly, $b_n = 1$ for all n. Let $a_n = a_{Nn}/a_{Dn}$ for $n = 2, 3, 4, \ldots$ The sequence $a_{Nn} = 9, 16, 25, \ldots$ is generated by

$$a_{Nn} = a_{N(n-1)} + (2n+1) \tag{73}$$

while the sequence $a_{Dn} = 35, 63, 99, \ldots$ is generated by

$$a_{Dn} = a_{D(n-1)} + 4(2n+1) \tag{74}$$

with $a_{N1} = 4$ and $a_{D1} = 15$. To evaluate ξ using equations (72) , use

$$\begin{aligned} a_1 &= \frac{a_1}{b_1} = \frac{1}{5} \\ b_n &= 1 \qquad\qquad\qquad\qquad \text{for } n = 2, 3, 4, \ldots \\ a_n &= -\left(\frac{a_{N(n-1)} + (2n+1)}{a_{D(n-1)} + 4(2n+1)}\right)x \end{aligned}$$

4.2 Improving the Convergence of the Continued Fraction

The range of convergence for the continued fraction representing ξ is $-1 \leq x < \infty$. The efficiency of the iteration method used to evaluate ξ is measured by the number of continued fraction levels required and is greatly dependent on the value of x being considered. The iteration converges most rapidly for small values of x, near $x = 1$. The number of levels required in the evaluation increases as x gets larger. The convergence properties can be improved for large x by relating the original $\xi(x)$ to $\xi(\eta)$ where

$$\eta = \frac{\sqrt{1+x} - 1}{\sqrt{1+x} + 1} \tag{75}$$

Note that η approaches $+1$ as x approaches ∞.

Recall that $F(\frac{1}{2}, 1; \frac{3}{2}; -x) = (\tan^{-1}\sqrt{x})/\sqrt{x}$ from which

$$\frac{\frac{1}{2}\psi}{\tan\frac{1}{2}\psi} = F(\tfrac{1}{2}, 1; \tfrac{3}{2}; -\tan^2\tfrac{1}{2}\psi) \qquad \text{and} \qquad \frac{\frac{1}{4}\psi}{\tan\frac{1}{4}\psi} = F(\tfrac{1}{2}, 1; \tfrac{3}{2}; -\tan^2\tfrac{1}{4}\psi)$$

Isolating ψ gives

$$\psi = 2\tan\tfrac{1}{2}\psi F(\tfrac{1}{2}, 1; \tfrac{3}{2}; -\tan^2\tfrac{1}{2}\psi) = 4\tan\tfrac{1}{4}\psi F(\tfrac{1}{2}, 1; \tfrac{3}{2}; -\tan^2\tfrac{1}{4}\psi).$$

Therefore

$$F(\tfrac{1}{2}, 1; \tfrac{3}{2}; -\tan^2\tfrac{1}{2}\psi) = \frac{2\tan\frac{1}{4}\psi}{\tan\frac{1}{2}\psi} F(\tfrac{1}{2}, 1; \tfrac{3}{2}; -\tan^2\tfrac{1}{4}\psi) \tag{76}$$

From trigonometric identities,

$$\frac{\tan\frac{1}{4}\psi}{\tan\frac{1}{2}\psi} = \frac{1}{\sec\frac{1}{2}\psi + 1} \qquad \text{and} \qquad \tan^2\tfrac{1}{4}\psi = \frac{\sec\frac{1}{2}\psi - 1}{\sec\frac{1}{2}\psi + 1}$$

or in terms of $x = \tan^2 \frac{1}{2}\psi$,

$$\frac{\tan \frac{1}{4}\psi}{\tan \frac{1}{2}\psi} = \frac{1}{\sqrt{1+x}+1} \qquad \text{and} \qquad \tan^2 \tfrac{1}{4}\psi = \frac{\sqrt{1+x}-1}{\sqrt{1+x}+1} = \eta.$$

Denoting $F(x) = F(\frac{1}{2}, 1; \frac{3}{2}; -\tan^2 \frac{1}{2}\psi)$ and $F(\eta) = F(\frac{1}{2}, 1; \frac{3}{2}; -\tan^2 \frac{1}{4}\psi)$, equation (76) becomes

$$F(x) = \frac{2}{\sqrt{1+x}+1} F(\eta). \tag{77}$$

Since $G(x)$ is defined by $F(x) = 1/(1 + xG(x))$, equation (77) can be written as

$$\frac{1}{1+xG(x)} = \frac{2}{\sqrt{1+x}+1}\left(\frac{1}{1+\eta G(\eta)}\right).$$

Rearranging,

$$2(1 + xG(x)) = (\sqrt{1+x}+1)(1+\eta G(\eta)).$$

Isolating $G(x)$ in the above equation and using $(\sqrt{1+x}-1)(\sqrt{1+x}+1) = x$ yields

$$G(x) = \frac{1}{2(\sqrt{1+x}+1)}(1 + G(\eta)). \tag{78}$$

But $\xi(x)$ is defined by $G(x) = 1/(3 + 4x\xi(x))$, so that

$$\begin{aligned}\frac{1}{3+4x\xi(x)} &= \frac{1}{2(\sqrt{1+x}+1)}\left(1 + \frac{1}{3+4\eta\xi(\eta)}\right) \\ &= \frac{2}{\sqrt{1+x}+1}\left(\frac{1+\eta\xi(\eta)}{3+4\eta\xi(\eta)}\right)\end{aligned}$$

Rearranging gives

$$6 + 8x\xi(x) = (\sqrt{1+x}+1)\left(3 + \frac{\eta\xi(\eta)}{1+\eta\xi(\eta)}\right).$$

Finally, isolating $\xi(x)$ yields

$$\xi(x) = \frac{1}{8(\sqrt{1+x}+1)}\left(3 + \frac{\xi(\eta)}{1+\eta\xi(\eta)}\right). \tag{79}$$

To evaluate $\xi(x)$, first compute η from equation (75) and obtain $\xi(\eta)$ using the continued fraction (61). Then equation (79) gives $\xi(x)$ from x, η, and $\xi(\eta)$.

Table 1 is a comparison of the average number of continued fraction levels required to compute $\xi(x)$ to eight significant digits directly versus using the intermediate result $\xi(\eta)$ for various values of λ and T. A substantial improvement in convergence is seen throughout the entire table when the continued fraction is evaluated using η. The direct method of computing ξ always requires a greater number of levels, sometimes exceeding 150, to obtain ξ to the desired accuracy. The right half of

Table 1. Number of continued fraction levels needed to compute ξ to eight significant digits directly vs. using $\xi(\eta)$ ($\xi(z)$, $\xi(\eta)$).

	T=0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	3.0	5.0	7.0	9.0	11.0	13.0	15.0	17.0	19.0
λ=-0.9	†,33	118,24	81,19	62,17	50,15	42,14	36,13	31,12	27,11	24,10	12,6	46,9	81,9	105,10	126,10	139,10	†,10	†,10	†,10
-0.8	†,30	102,22	70,18	53,16	43,14	36,13	30,11	26,11	23,10	20,9	15,7	35,8	50,9	61,9	70,9	77,9	83,9	89,9	94,9
-0.7	†,28	88,20	60,17	46,14	37,13	30,11	26,10	22,10	19,9	17,8	16,7	29,8	38,9	45,9	50,9	55,9	59,9	63,9	66,9
-0.6	142,26	75,19	51,15	39,13	31,12	25,10	21,9	18,9	16,8	14,7	16,7	25,8	31,8	36,9	40,9	43,9	46,9	49,9	51,9
-0.5	121,24	64,17	44,14	33,12	26,11	21,9	18,9	15,8	13,7	11,6	15,7	22,8	26,8	30,8	33,8	36,9	38,9	40,9	42,9
-0.4	102,22	54,16	37,13	28,11	22,10	18,9	15,8	13,7	11,6	10,6	14,7	19,7	23,8	26,8	29,8	31,8	33,8	35,8	36,9
-0.3	85,20	45,14	31,12	23,10	19,9	15,8	13,7	11,6	10,6	8,5	12,6	17,7	20,7	23,8	25,8	27,8	29,8	30,8	32,8
-0.2	60,18	36,13	25,10	19,9	16,8	13,7	11,7	10,6	9,6	8,5	12,6	16,7	18,7	21,8	23,8	24,8	26,8	27,8	28,8
-0.1	50,15	27,11	20,9	16,8	13,7	11,7	10,6	9,6	8,5	7,5	11,6	14,7	17,7	19,7	20,7	22,8	23,8	24,8	25,8
0.	25,10	18,9	15,8	13,7	11,6	10,6	9,6	8,5	7,5	6,5	10,6	13,6	15,7	17,7	19,7	20,7	21,8	22,8	23,8
0.1	13,7	12,7	11,6	10,6	9,6	9,5	8,5	7,5	7,5	6,4	9,5	12,6	14,7	16,7	17,7	18,7	19,7	20,7	21,8
0.2	9,6	9,6	9,5	8,5	8,5	7,5	7,5	6,5	6,4	6,4	9,5	11,6	13,6	14,7	16,7	17,7	18,7	19,7	19,7
0.3	7,5	7,5	7,5	7,5	7,5	6,4	6,4	6,4	5,4	5,4	8,5	10,6	12,6	13,6	14,7	15,7	16,7	17,7	18,7
0.4	6,4	6,4	6,4	6,4	6,4	6,4	5,4	5,4	5,4	4,3	7,5	10,6	11,6	12,6	13,6	14,7	15,7	16,7	17,7
0.5	5,4	5,4	5,4	5,4	5,4	5,4	5,4	4,3	4,3	4,3	7,5	9,5	11,6	12,6	13,6	14,7	15,7	15,7	15,7
0.6	5,4	5,4	4,4	4,3	4,3	4,3	4,3	4,3	4,3	3,3	7,5	8,5	10,6	11,6	12,6	13,6	13,6	14,7	14,7
0.7	4,3	4,3	4,3	4,3	4,3	4,3	3,3	3,3	3,2	3,3	6,4	8,5	9,5	10,6	11,6	12,6	12,6	13,6	14,7
0.8	3,3	3,3	3,3	3,3	3,3	3,2	3,2	3,3	3,3	4,3	6,4	7,5	8,5	9,6	10,6	11,6	12,6	12,6	13,6
0.9	3,2	3,2	2,2	2,2	3,2	3,3	3,3	3,3	3,3	4,3	6,4	7,5	8,5	9,5	10,6	10,6	11,6	11,6	12,6

†Over 150 levels required to compute $\xi(z)$.

Entries to the left of the dotted line represent hyperbolic orbits; those to the right of the dotted line represent elliptic orbits.

the table, where T is between 1 and 20, represents larger values of x. Computing ξ using η requires 10 or fewer levels in this region while computing $\xi(x)$ directly often requires over 20. Both methods take a larger number of levels to converge in the upper left corner of the table since x and η are both approaching -1, the lower limit of convergence of the continued fraction. Although some extra manipulations are necessary to obtain $\xi(x)$ from $\xi(\eta)$, this method is significantly more efficient due to the computational saving in evaluating the continued fraction.

4.3 Solving the Cubic Equation

The cubic equation which must be solved for y using the new method is

$$y^3 - (1 + h_1)y^2 - h_2 = 0. \tag{58}$$

This can be interpreted graphically using the transformation $y = (1 + h_1)w$. Substituting this expression for y, the cubic equation becomes

$$w^3 - w^2 = w^2(w - 1) = \frac{h_2}{(1 + h_1)^3}.$$

A plot of w versus $h_2/(1+h_1)^3$ is shown in Fig. 7. Compare this graph to Fig. 3 which shows a similar plot of the cubic in Gauss' method. There are multiple real roots for w, and hence y, only in the small region $-\frac{4}{27} < h_2/(1 + h_1)^3 < 0$. For most problems of practical interest, the quantity $h_2/(1 + h_1)^3$ is positive and there is only one real root for w. This corresponds to a value of $y > 1 + h_1$. If there are three real roots of the cubic, the value of y is chosen so that w lies along the upper branch of the curve, $w > \frac{2}{3}$. This choice represents the maximum of the three roots for y. Note that there is no need to distinguish cases of transfer angle less than or greater than 180 degrees. In all cases, the desired root lies along the upper branch of the cubic. The point $w = 1$ where the curve intersects the w-axis represents the case of a parabolic orbit with $\lambda = 0$.

Analytic Solution of the Cubic Equation

The cubic equation can be solved analytically using the transformation

$$y = \frac{(1 + h_1) - 2|1 + h_1|z}{3} \tag{80}$$

After some algebraic manipulations the cubic becomes

$$4z^3 - 3z = b \tag{81}$$

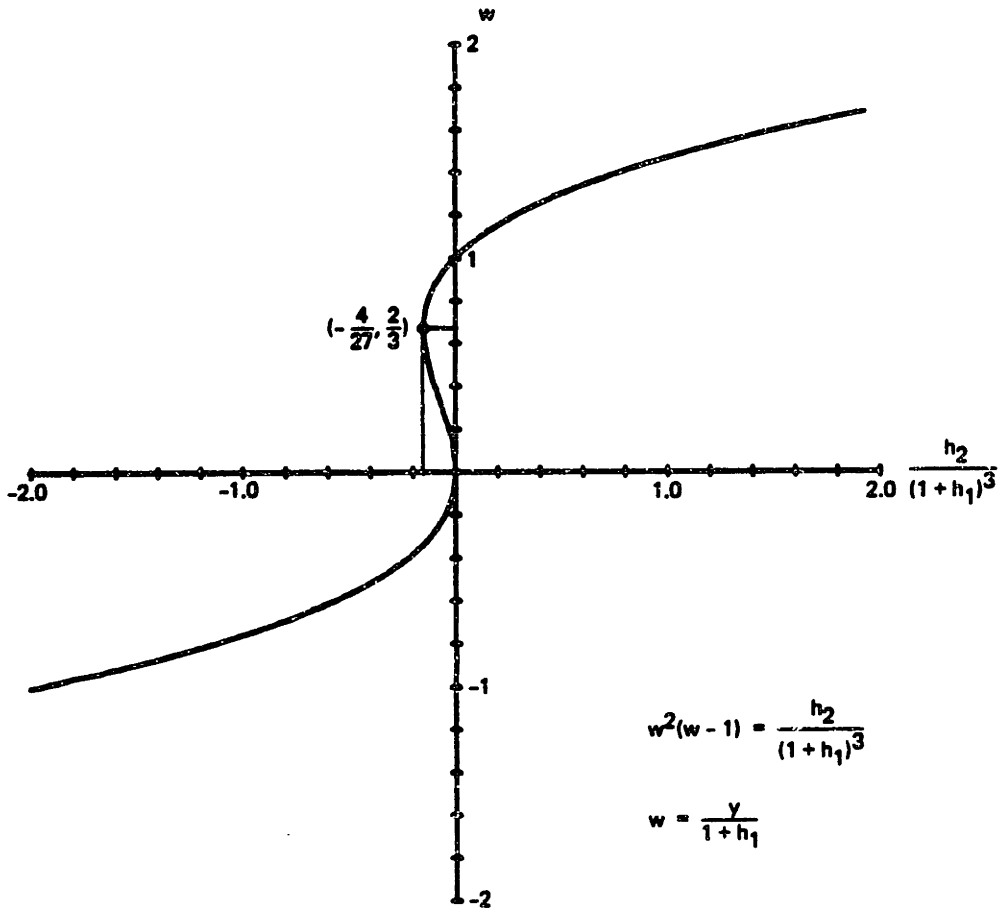


Figure 7. Cubic Equation of New Method

where

$$b = \frac{-\left(27h_2 + 2(1+h_1)^3\right)}{2|1+h_1|^3} \tag{82}$$

First, consider cases where b is non-negative. When $b \geq 1$, a form of Cardan's solution for cubic equations is applied. The single real root for z is found from

$$z = \frac{1}{2}(u + \frac{1}{u}) \qquad \text{where} \qquad u = \left(b + \sqrt{b^2 - 1}\right)^{\frac{1}{3}}.$$

When $0 \leq b < 1$, there are three real roots for z. If the single positive root is denoted $z_1 = \cos\gamma$, then

$$4z_1^3 - 3z_1 = 4\cos^3\gamma - 3\cos\gamma = b.$$

From trigonometry $\cos^3\gamma - 3\cos\gamma = \cos 3\gamma$, so that $\gamma = \frac{1}{3}\cos^{-1} b$. The two negative roots for z are found from $z_2 = -\cos(\frac{\pi}{3} + \gamma)$ and $z_3 = -\cos(\frac{\pi}{3} - \gamma)$. Cases of negative b are handled similarly by setting $z = -Z$. The cubic equation (81) then becomes

$$4Z^3 - Z = -b$$

. Thus for $-b \geq 1$, or $b \leq 1$, the sole real root is $Z = \frac{1}{2}(u + \frac{1}{u})$ where $u = \left(-b + \sqrt{(-b)^2 - 1}\right)^{\frac{1}{3}}$. For $0 < -b < 1$, or $-1 < b < 0$, the three real roots are $Z_1 = \cos\gamma$, $Z_2 = -\cos(\frac{\pi}{3} + \gamma)$, and $Z_3 = -\cos(\frac{\pi}{3} - \gamma)$ where $\gamma = \frac{1}{3}\cos^{-1}(-b)$. Once all the real roots for z are obtained, the corresponding values of y are computed from equation (80).[16] If there are multiple real roots, the maximum value is chosen for y.

Another analytic solution of the cubic is possible if $h_2/(1+h_1)^3$ is positive. There is only one real root for y in this case. Substitute the following transformation equation

$$y = \sqrt{\frac{3h_2}{1+h_1}}\frac{1}{z} \tag{83}$$

into the cubic equation to obtain

$$\left(\frac{3h_2}{1+h_1}\right)^{\frac{3}{2}}\frac{1}{z^3} - \frac{3h_2}{z^2} = h_2.$$

Multiplying by $\left((1+h_1)/3h_2\right)^{\frac{3}{2}}$ gives

$$\frac{3}{z^3}\left(1 - z\sqrt{\frac{(1+h_1)^3}{3h_2}}\right) = \sqrt{\frac{(1+h_1)^3}{3h_2}}.$$

Rewriting as a cubic equation in z yields

$$\frac{1}{3}z^3 + z = \sqrt{\frac{3h_2}{(1+h_1)^3}} \tag{84}$$

Cardan's solution can be applied directly to this cubic equation. If b is defined as

$$b = \frac{3}{2}\sqrt{\frac{3h_2}{(1+h_1)^3}} = \sqrt{\frac{27h_2}{4(1+h_1)^3}} \tag{85}$$

then

$$\begin{aligned} z &= \left(\sqrt{b^2+1}+b\right)^{\frac{1}{3}} - \left(\sqrt{b^2+1}-b\right)^{\frac{1}{3}} \\ &= \left(\sqrt{b^2+1}+b\right)^{\frac{1}{3}} - \frac{1}{\left(\sqrt{b^2+1}+b\right)^{\frac{1}{3}}} \end{aligned} \tag{86}$$

An improved algebraic expression for z is obtained using $a = \left(\sqrt{b^2+1}+b\right)^{\frac{2}{3}}$ from which

$$z = \frac{2ab}{1+a+a^2} \tag{87}$$

It can be seen from the defining equations for z and b that the solution of the cubic obtained from this method lies along the upper branch of the curve in Fig. 7. The final value of y is found from equation (83) after z is computed using equation (86) or (87). This method can be employed in most problems of practical interest; the restriction on $h_2/(1+h_1)^3$ excludes only cases of large transfer angle, θ approaching 2π, and small flight times.

Newton-Rapheson Iteration

An algebraically simpler procedure for solving the cubic is the well-known Newton-Rapheson iteration method. An initial guess for y, y_o, is chosen. Successive values of y are generated from

$$y_{n+1} = y_n - \frac{y_n^3 - (1+h_1)y_n^2 - h_2}{3y_n^2 - 2(1+h_1)y_n}. \tag{88}$$

The numerator of the second term on the right is the value of the cubic at y_n while the denominator is the first derivative of the cubic with respect to y at y_n. The iteration is complete when $y_n - y_{n-1}$ is within some specified tolerance. This iteration for y must be performed during each step of the overall iteration for z. On the first iteration for z, the choice of y_o depends on the values of the coefficients h_1 and h_2. Referring to Fig. 7, a value of $y_o = \frac{2}{3}(1+h_1)$ is chosen if $h_2/(1+h_1)^3 \geq -\frac{4}{27}$ while $y_o = 0$ is selected when $h_2/(1+h_1)^3 < -\frac{4}{27}$. On each subsequent interation for z, the initial value y_o depends on the value of y from the previous iteration. If the last value of y is negative and the the new values of h_1 and h_2 are such that $h_2/(1+h_1)^3 > -\frac{4}{27}$, then y_o is again chosen to be $\frac{2}{3}(1+h_1)$. This is done to avoid using the Newton-Rapheson technique in the region where the derivative of the cubic changes value rapidly. The new y_o is set equal to the previous value of y if that value is positive. This greatly reduces the number of iterations required to obtain y.

Solution using a Continued Fraction

A third method of solving the cubic equation can be developed which employs a continued fraction. First, consider the transformation

$$y = \frac{2}{3}(1+h_1)\left(\frac{b}{z}+1\right) \tag{89}$$

where

$$b = \sqrt{\frac{27h_2}{4(1+h_1)^3}+1} \tag{90}$$

Substituting this expression for y in the cubic equation, there results

$$\frac{8}{27}(1+h_1)^3\left(\frac{b}{z}+1\right)^3 - \frac{4}{9}(1+h_1)^3\left(\frac{b}{z}+1\right)^2 - h_2 = 0.$$

Separating terms containing z from the constant terms gives

$$\left(\frac{b}{z}+1\right)^2\left(2\left(\frac{b}{z}+1\right)-3\right) = \frac{27h_2}{4(1+h_1)^3}.$$

From the definition of b the term on the right is seen to be b^2-1. Now in terms of b and z the cubic equation becomes

$$\left(\frac{b}{z}+1\right)^2\left(2\frac{b}{z}-1\right) = b^2-1$$

or, rearranging

$$z^3 - 3z = 2b \tag{91}$$

If the value of b is confined to the range $0 < b < 1$, trigonometric functions can be substituted for z and b in the cubic. Specifically, if $z = 2\cos\frac{1}{3}\gamma$, then

$$4\cos^3\tfrac{1}{3}\gamma - 3\cos\tfrac{1}{3}\gamma = \cos\gamma = b$$

Using trigonometric identities to relate $\cos\frac{1}{3}\gamma$ and $\cos\gamma$ to $\sin\frac{1}{2}\gamma$, b and z become

$$z = 2\cos\tfrac{1}{3}\gamma = 2(1-2\sin^2\tfrac{1}{6}\gamma) \qquad \text{and} \qquad \sin^2\tfrac{1}{2}\gamma = \tfrac{1}{2}(1-\cos\gamma) = \tfrac{1}{2}(1-b).$$

The expression for z can be written as

$$z = 2\left(1-(1-b)\left(\frac{\sin^2\frac{1}{6}\gamma}{\sin^2\frac{1}{2}\gamma}\right)\right)$$

by multiplying the second term in parentheses by $\frac{1}{2}(1-b)/\sin^2\frac{1}{2}\gamma = 1$. Let $K = \sin\frac{1}{6}\gamma/\sin\frac{1}{2}\gamma$. Then K can be expressed as the ratio of two hypergeometric functions of $\sin^2\frac{1}{2}\gamma$.

$$K = \frac{1}{3}\frac{F(\frac{2}{3},\frac{4}{3};\frac{3}{2};\sin^2\frac{1}{2}\gamma)}{F(\frac{2}{3},\frac{1}{3};\frac{1}{2};\sin^2\frac{1}{2}\gamma)}$$

This method can be extended to include values of $b \geq 1$ by using hyperbolic functions. In these cases,

$$z = 2\cosh\tfrac{1}{3}\gamma = 2\left(1-(1-b)\left(\frac{\sinh^2\frac{1}{6}\gamma}{\sinh^2\frac{1}{2}\gamma}\right)\right),$$

$$b = \cosh\gamma,$$

and

$$K = \frac{\sinh\frac{1}{6}\gamma}{\sinh\frac{1}{2}\gamma} = \frac{1}{3}\frac{F(\frac{2}{3},\frac{4}{3};\frac{3}{2};-\sinh^2\frac{1}{2}\gamma)}{F(\frac{2}{3},\frac{1}{3};\frac{1}{2};-\sinh^2\frac{1}{2}\gamma)}.$$

Since $-\sinh^2\frac{1}{2}\gamma = \frac{1}{2}(1-b)$, K is a function of

$$u = \tfrac{1}{2}(1-b) = \begin{cases} \sin^2\frac{1}{2}\gamma & 0 \leq b < 1 \\ -\sinh^2\frac{1}{2}\gamma & b \geq 1 \end{cases}$$

K can be expanded into a continued fraction as shown below[17]

$$K(u) = \frac{1}{3}\cfrac{1}{3-\cfrac{4u}{9-\cfrac{40u}{15-\cfrac{70u}{21-\cdots}}}} = \cfrac{\frac{1}{3}}{1-\cfrac{\frac{4}{27}u}{1-\cfrac{\frac{40}{135}u}{1-\cfrac{\frac{70}{315}u}{1-\cdots}}}} \tag{92}$$

Also the final expression for z is

$$z = 2\big(1-(1-b)K^2(u)\big). \tag{93}$$

To solve the cubic, first compute b from the coefficients h_1 and h_2 using equation (90). Next, evaluate the continued fraction for $K(u)$ using $u = \frac{1}{2}(1-b)$. If $K(u)$ is written as

$$K(u) = \cfrac{\frac{1}{3}}{1-\cfrac{k_2u}{1-\cfrac{k_3u}{1-\cfrac{k_4u}{1-\cdots}}}}$$

then the coefficients $k_n = k_{Nn}/k_{Dn}$ for $n = 2,3,4,\ldots$ can be related to the k_{n-1} coefficients. The numerator sequence $k_{Nn} = 4, 40, 70, \ldots$ is generated from

$$k_{Nn} = \begin{cases} k_{N(n-1)} + 6(2n-1) & n = 2,4,6,\ldots \\ k_{N(n-1)} + 12(2n-1) & n = 3,5,7,\ldots \end{cases}$$

with $k_{N1} = -2$. The denominator sequence $k_{Dn} = 27, 135, 315, \ldots$ is generated from

$$k_{Dn} = k_{D(n-1)} + 36(2n-1) \qquad n = 2,3,4,\ldots$$

with $k_{D1} = -9$. Thus, to evaluate $K(u)$ with the top down method, use

$$b_n = 1 \qquad n = 1,2,3,\ldots$$

$$a_n = k_n u = \begin{cases} \dfrac{k_{N(n-1)} + 6(2n-1)}{k_{D(n-1)} + 36(2n-1)} u & n = 2, 4, 6, \ldots \\ \dfrac{k_{N(n-1)} + 12(2n-1)}{k_{D(n-1)} + 36(2n-1)} u & n = 3, 5, 7, \ldots \end{cases}$$

and $u_1 = a_1/b_1 = \frac{1}{3}$ in equations (72). Having computed $K(u)$, the value of x is obtained from equation (93) and y is computed using b and x in equation (89).

Although this method requires a significant amount of computation, it has the advantage that it does not involve a choice among multiple roots for y. The desired value of y is generated automatically. Since b must be a real number, the solution is restricted to areas where $h_2/(1+h_1)^3 \geq -\frac{4}{27}$. When $0 \leq b < 1$, $-\frac{4}{27} \leq h_2/(1+h_1)^3 < 0$ and when $b \geq 1$, $h_2/(1+h_1)^3 \geq 0$. Therefore, referring to Fig. 7, it can be seen that this method gives the values of y which lie along the upper branch of the cubic. A minor disadvantage of this method is that it cannot be applied in cases where $h_2/(1+h_1)^3 < -\frac{4}{27}$. While no final solutions for y lie in this region, some intermediate steps in the overall iteration may fall in this region.

Chapter 5

Comparison of Gauss' Method and the New Method

5.1 Results

WHILE THE DERIVATIONS of Gauss' method and the new method for solving Lambert's problem are different, the final equations are very similar. Figure 8 lists the key equations for each method. Both involve the simultaneous solution of two equations, one a cubic in y, for x and y. However, x and y are defined differently for each method. In Gauss' method, elliptic orbits correspond to $0 < x \leq 1$ and hyperbolic orbits correspond to $-\infty \leq x < 0$. For the new method, elliptic orbits occur in the range $0 < x < \infty$ while the range $-1 \leq x < 0$ includes all hyperbolic orbits. Parabolic orbits are represented by $x = 0$ in both methods. The variable x in the new method is simply related to the quantity x in Gauss' method using trigonometric or hyperbolic function identities.

Figures 9 and 10 show flowcharts of the successive substitution procedure for Gauss' method and the new method respectively. The dimensionless parameters λ and T are considered as inputs. For any specified λ and T, the new method requires more algebra for each iteration since there are two coefficients to be found for the cubic equation and equation (44) is quadratic in x. There is no need to test the value of λ as in Gauss' method since the new equations are valid for $-1 < \lambda \leq 1$. The efficiency of either procedure is measured by the number of iterations necessary to compute x to a given accuracy.

Tables 2 and 3 show the number of iterations required to compute x using Gauss' method to eight and twelve significant digits respectively. The initial value of x used to generate these tables was

$$x_o = \begin{cases} 0 & \text{parabola, hyperbola} \\ \dfrac{(1-\lambda)^2}{2(1+\lambda^2)} = \dfrac{l}{1+2l} & \text{ellipse} \end{cases}$$

Cases of elliptic and hyperbolic orbits are easily distinguished by comparing the value of m to that of the parabolic orbit given by equation (16). The initial value of x for elliptic orbits corresponds to a circular orbit. Note the rapid convergence of Gauss' method in the lower left corner of the tables. In this region, x is near 0 and the transfer angle is small. Gauss' method was designed for problems of this type. The quantity ξ was defined to be of order x^2 so that it is small for small x. Indeed,

Gauss' Method

$$l = \frac{(1-\lambda)^2}{4\lambda} \qquad m = \frac{T^2}{64\lambda^3}$$

$$x = \begin{cases} \sin^2 \frac{1}{4}(E_2 - E_1) & -\infty < x \leq 1 \\ -\sinh^2 \frac{1}{4}(H_2 - H_1) & \end{cases}$$

$$h = \frac{m}{\frac{5}{6} + l + \xi}$$

$$\xi = \cfrac{\frac{2}{35}x^2}{1 + \frac{2}{35}x - \cfrac{\frac{40}{63}x}{1 - \cfrac{\frac{4}{99}x}{1 - \cdots}}}$$

$$y^3 - y^2 - hy - \tfrac{1}{9}h = 0$$

$$\frac{m}{y^2} = x + l$$

New Method

$$l = \left(\frac{1-\lambda}{1+\lambda}\right)^2 \qquad m = \frac{T^2}{(1+\lambda)^6}$$

$$x = \begin{cases} \tan^2 \frac{1}{4}(E_2 - E_1) & -1 \leq x < \infty \\ -\tanh^2 \frac{1}{4}(H_2 - H_1) & \end{cases}$$

$$h_1 = \frac{(l+x)^2(1+\xi(1+3x))}{(1+2x+l)(1+x(4\xi+1))}$$

$$h_2 = \frac{m(1+\xi(x-l))}{(1+2x+l)(1+x(4\xi+1))}$$

$$\xi = \cfrac{\frac{1}{5}}{1 + \cfrac{\frac{9}{35}x}{1 + \cfrac{\frac{16}{63}x}{1 + \cdots}}}$$

$$y^3 - (1+h_1)y^2 - h_2 = 0$$

$$\frac{m}{y^2} = (1+x)(l+x)$$

Figure 8. Key equations of Gauss' Method and the New Method

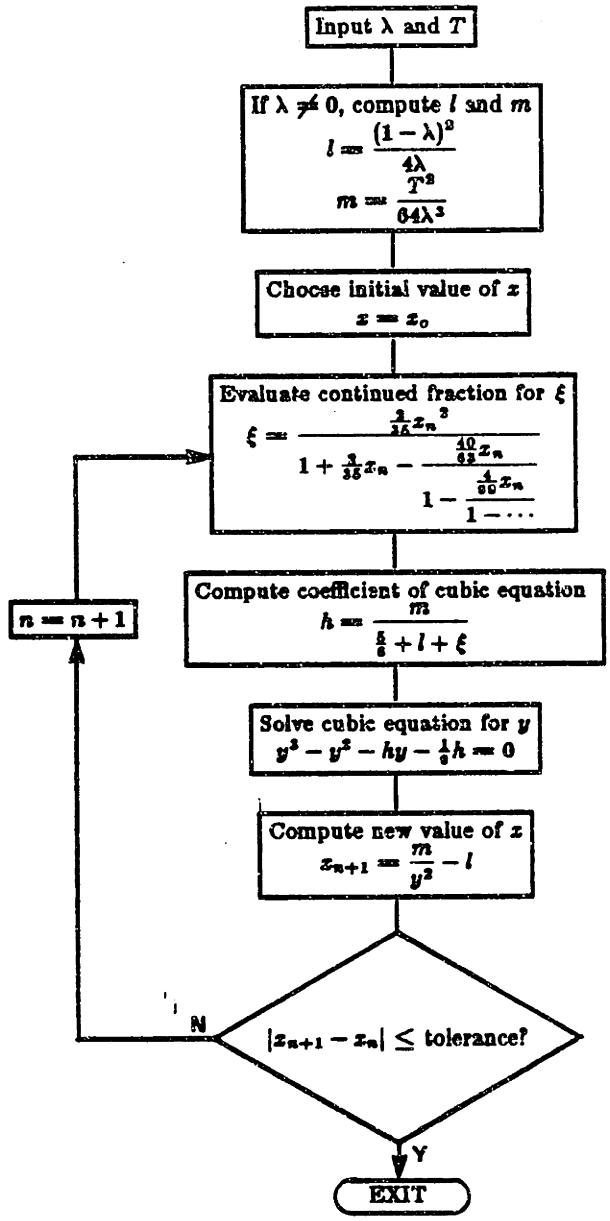


Figure 9. Flowchart of the Iteration Process using Gauss' Method

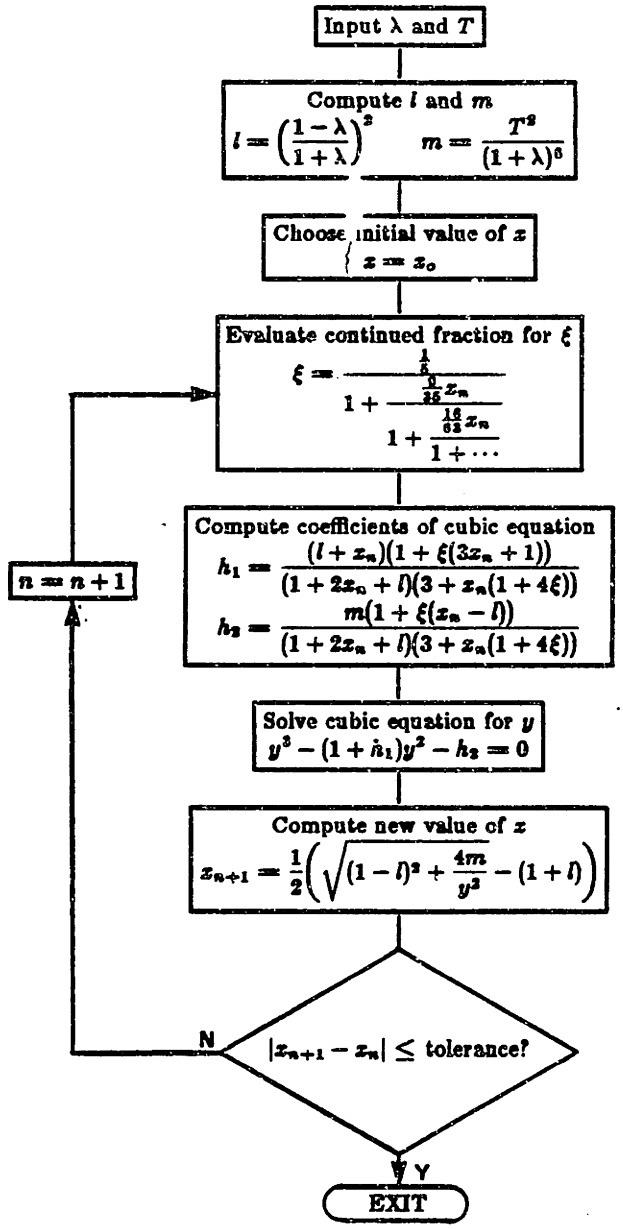


Figure 10. Flowchart of the Iteration Process using the New Method

$\xi = 0$ when $x = 0$. The coefficient of the cubic, h, is nearly independent of ξ for small x so that y is nearly constant.

A major disadvantage of Gauss' method is the singularity at $\lambda = 0$ shown in Tables 2 and 3. This follows immediately from the equations for m and l, which are undefined at $\lambda = 0$. This corresponds to the case of 180-degree transfer angle, $\theta = \pi$. Another disadvantage is the non-uniform behavior over the range of λ and T considered. While convergence is rapid for λ near $+1$ and small T, there is an increase in the number of iterations for cases with positive λ and larger values of T. These are cases where x, and hence ξ, is becoming significantly greater than 0 and no longer has negligible effect on h. When λ is negative, there is a substantial increase in the number of iterations. Gauss' method converges for all elliptic cases considered but sometimes requires over 100 iterations. For most hyperbolic cases, Gauss' method fails to converge. This is due to difficulties in solving the cubic equation discussed in Chapter 2. In comparing Table 3 with Table 2, it is seen that only one or two more iterations are needed to obtain the next four significant digits in the lower left corner of the table where x is small. The difference becomes greater for positive λ as T increases and for all negative λ.

The new method was designed to converge rapidly for any case independent of the value of x. The near uniform convergence behavior of the new method is seen in Tables 4 and 5; these tables show the number of iterations needed to compute x to eight and twelve significant figures using the new method for the same range of λ and T considered for Gauss' method. The initial value chosen for x was identical to that used for Gauss' method. In particular,

$$x_o = \begin{cases} 0 & \text{parabola, hyperbola} \\ l & \text{ellipse} \end{cases}$$

where $x_o = l$ corresponds to a circular orbit. Ellipses and hyperbolas were distinguished by comparing m to that of the parabolic orbit, equation (65). The singularity at $\lambda = 0$ has been removed from the new method by redefining m and l. The rapid convergence in the lower left corner of the tables noted using Gauss' method is retained using the new method. Note also that there is no significant difference in convergence for positive or negative values of λ. Perhaps the most important advantage of the new method is revealed by comparing the difference in the number of iteration steps needed to obtain twelve versus eight significant digits. As Tables 4 and 5 show, only one more iteration is necessary to obtain four more significant figures in all cases.

The definitions of m and l for the new method show that there is a singularity at $\lambda = -1$. This corresponds to a 360-degree transfer angle, $\theta = 2\pi$, which is not of great practical interest. It is instructive, however, to investigate the behavior of the new method for values of λ approaching -1.

Tables 6 and 7 show the number of iterations required for convergence to eight and twelve significant figures in cases where λ varies between -0.90 and -0.99. The uniform behavior extends to these cases except for a narrow region near $T = 5$. The increase in the number of iteration steps in this region is not the result of a poor initial guess for x but rather an inherent property of the physical problem. The number of iteration steps approaches a maximum when the solution of the problem approaches the minimum energy orbit for a given value of λ. The minimum energy orbit connecting two position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ is defined to be the orbit having semi-major axis

$$a_m = \tfrac{1}{3}s = \tfrac{1}{4}(r_1 + r_2 + c)$$

Table 8 shows the values of T as well as the number of iterations required for convergence for minimum energy orbits where λ varies from -0.90 to -0.99. Inspecting the final values of y, h_1, and h_2 in these cases reveals that the solution of the cubic approaches the point of transition between one negative real root and three real roots ($w = \frac{2}{3}$ in Fig. 7).

Difficulties have also been found near the minimum energy orbit when using a Newton-Rapheson technique to solve Lambert's problem. In the formulation given by Lancaster and Blanchard[18,19] and Battin[6], the quantity T expressed as a function of the chosen iterated variable has a change in curvature in this region when λ is near -1. This can cause the iteration process to diverge. Lancaster and Blanchard recommend changing from the Newton-Rapheson method to a regula falsi technique in this area. Although it performed less efficiently, the new method successfully converged in all test cases both near and at the minimum energy orbit with no change in the procedure or choice of initial value for the iterated variable.

Table 2. Number of iterations required to compute x to eight significant digits using Gauss' method.

	T=0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	3.0	5.0	7.0	9.0	11.0	13.0	15.0	17.0	19.0
λ=-0.9	†	†	†	†	†	†	†	†	†	†	67	210	190	167	151	134	95	132	141
-0.8	†	†	†	†	†	†	†	†	†	†	60	88	85	81	73	68	78	82	85
-0.7	†	†	†	†	†	†	†	†	†	†	44	52	52	49	48	54	57	59	61
-0.6	†	†	†	†	†	†	†	†	†	†	32	35	35	34	39	42	43	45	46
-0.5	†	†	†	†	†	†	†	†	†	†	24	26	19	29	32	34	35	37	38
-0.4	†	†	†	†	†	†	†	†	†	†	18	20	22	24	27	28	30	30	32
-0.3	†	†	†	†	†	†	†	†	29	17	14	14	19	21	23	24	25	26	26
-0.2	†	†	†	†	†	76	25	17	13	10	12	13	17	18	20	21	22	22	23
-0.1	†	†	†	27	16	12	10	10	9	8	10	12	14	15	17	18	19	19	21
0.	S I N G U L A R I T Y																		
0.1	4	4	5	5	6	6	6	5	5	5	7	9	11	13	13	14	16	16	17
0.2	3	4	4	4	5	5	5	5	4	4	7	9	10	11	13	13	13	14	15
0.3	3	3	3	4	4	4	4	4	4	4	6	8	10	10	11	12	12	13	14
0.4	3	3	3	3	3	3	3	3	3	3	5	7	9	10	10	11	12	12	13
0.5	2	3	3	3	3	3	3	3	3	3	5	6	8	8	9	10	10	11	11
0.6	2	3	3	3	3	3	3	3	3	3	5	7	7	7	8	9	9	10	11
0.7	2	2	3	3	3	3	3	3	3	3	4	5	7	7	8	8	9	10	10
0.8	2	2	2	2	2	2	2	3	3	3	4	5	6	6	7	8	8	8	9
0.9	2	2	2	2	2	2	2	2	2	3	4	5	6	6	7	8	7	8	8

†Gauss' method does not converge.

Entries to the left of the dotted line represent hyperbolic orbits; those to the right of the dotted line represent elliptic orbits.

Table 3. Number of iterations required to compute x to twelve significant digits using Gauss' method.

	T=0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	3.0	5.0	7.0	9.0	11.0	13.0	15.0	17.0	19.0
λ=-0.9	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	67	332	351	337	332	331	333	335	339
-0.8	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	78	143	151	154	158	161	164	168	171
-0.7	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	61	84	91	96	99	103	106	109	112
-0.6	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	45	58	63	67	70	73	77	79	82
-0.5	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	33	43	47	51	54	57	59	61	63
-0.4	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	27	33	38	41	44	46	48	50	51
-0.3	↑	↑	↑	↑	↑	↑	↑	↑	43	24	21	27	31	33	36	38	40	42	43
-0.2	↑	↑	↑	↑	↑	110	35	25	19	15	17	22	26	28	30	31	34	36	36
-0.1	↑	↑	↑	38	25	18	16	14	12	11	14	19	22	24	26	28	29	30	31
0.	S I N G U L A R I T Y																		
0.1	5	6	7	7	8	8	8	8	7	7	10	14	16	18	20	22	23	23	24
0.2	4	5	5	6	6	6	6	6	6	6	9	13	14	16	18	20	21	21	22
0.3	4	4	4	5	5	5	5	5	5	6	8	11	13	15	17	17	18	19	20
0.4	3	4	4	4	4	4	5	5	5	5	8	10	12	13	15	16	16	17	18
0.5	3	3	4	4	4	4	4	4	4	4	7	9	11	12	13	15	16	16	16
0.6	3	3	3	3	4	4	4	4	4	3	7	8	10	11	12	14	14	15	16
0.7	3	3	3	3	3	3	3	3	3	4	6	7	9	10	11	12	13	14	14
0.8	3	3	3	3	3	3	3	3	3	3	5	7	8	10	10	11	12	13	13
0.9	2	2	2	3	3	3	3	3	3	3	5	6	7	9	9	10	11	11	12

Entries to the left of the dotted line represent hyperbolic orbits; those to the right of the dotted line represent elliptic orbits.

Table 4. Number of iterations required to compute x to eight significant digits using the new method.

	T=0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	3.0	5.0	7.0	9.0	11.0	13.0	15.0	17.0	19.0
λ=-0.9	4	4	4	4	4	5	5	5	5	5	8	7	6	5	5	4	3	4	4
-0.8	4	4	4	4	4	4	5	5	5	5	7	5	5	4	4	4	4	4	4
-0.7	4	4	4	4	5	5	5	5	5	5	5	5	4	4	3	3	3	4	4
-0.6	4	4	4	4	4	5	5	5	5	5	5	4	3	3	4	4	4	4	4
-0.5	4	4	4	4	4	5	4	5	4	5	4	4	3	4	4	4	4	5	5
-0.4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5
-0.3	4	4	4	4	4	4	4	4	4	4	4	3	4	4	5	5	5	5	5
-0.2	3	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5
-0.1	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5
0.	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5
0.1	4	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5
0.2	3	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5
0.3	3	3	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5
0.4	3	3	4	4	4	4	4	4	3	4	4	5	5	5	5	5	5	5	5
0.5	3	3	3	3	3	3	3	3	3	3	4	5	5	5	5	5	5	5	5
0.6	3	3	3	3	3	3	3	3	3	3	4	5	5	5	5	5	5	5	5
0.7	3	3	3	3	3	3	3	3	3	3	4	4	5	5	5	5	5	5	5
0.8	3	3	3	3	3	3	3	3	3	3	4	4	5	5	5	5	5	5	5
0.9	2	3	2	3	2	3	3	3	3	3	4	4	5	5	5	5	5	5	5

Entries to the left of the dotted line represent hyperbolic orbits; those to the right of the dotted line represent elliptic orbits.

Table 5. Number of iterations required to compute x to twelve significant digits using the new method.

	T=0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	3.0	5.0	7.0	9.0	11.0	13.0	15.0	17.0	19.0
λ=-0.9	4	4	5	5	5	5	5	5	5	5	8	8	6	5	5	4	4	4	4
-0.8	5	5	5	5	5	5	5	5	5	5	7	6	5	5	4	4	4	4	4
-0.7	4	5	5	5	5	5	5	5	5	5	6	5	5	4	4	4	4	4	4
-0.6	4	5	5	5	5	5	5	5	5	5	5	4	4	4	4	5	5	5	5
-0.5	4	5	5	5	5	5	5	5	5	5	4	4	3	4	5	5	5	5	5
-0.4	4	5	5	5	5	5	5	5	5	5	4	4	4	5	5	5	5	5	5
-0.3	4	5	5	5	5	5	5	5	4	4	5	4	5	5	5	5	5	5	6
-0.2	4	4	4	4	4	4	4	4	4	4	5	4	5	5	5	5	6	6	6
-0.1	4	4	5	5	5	5	5	5	5	5	4	5	5	5	5	6	6	6	6
0.	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	6	6	6	6
0.1	4	5	5	5	5	5	5	5	5	4	4	5	5	5	6	6	6	6	6
0.2	4	4	4	4	5	5	4	4	4	4	4	5	5	5	6	6	6	6	6
0.3	4	4	4	4	4	4	4	4	4	4	5	5	5	5	6	6	6	6	6
0.4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	6	6	6	6	6
0.5	3	4	4	4	4	4	4	4	4	4	5	5	5	5	6	6	6	6	6
0.6	3	3	4	4	4	4	4	4	4	3	5	5	5	5	5	6	6	6	6
0.7	3	3	3	4	3	3	4	3	4	4	5	5	5	5	5	6	6	6	6
0.8	3	3	3	3	3	3	3	3	3	3	4	5	5	5	5	5	5	6	6
0.9	3	3	3	3	3	3	3	3	3	3	4	5	5	5	5	5	5	6	6

Entries to the left of the dotted line represent hyperbolic orbits; those to the right of the dotted line represent elliptic orbits.

Table 6. Number of iterations required to compute x to eight significant digits using the new method for λ near −1.

	T=0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	3.0	5.0	7.0	9.0	11.0	13.0	15.0	17.0	19.0
λ=−.99	4	4	4	4	4	4	5	5	5	5	10	14	8	6	5	5	4	3	4
−.98	4	4	4	4	4	4	5	5	5	5	9	12	8	6	5	5	4	3	4
−.97	4	4	4	4	4	4	4	4	5	5	9	11	7	6	5	4	4	3	4
−.96	4	4	4	4	4	4	5	5	5	5	9	10	7	6	5	4	4	3	4
−.95	4	4	4	4	4	5	5	5	5	5	9	9	7	5	5	4	4	4	4
−.94	4	4	4	4	4	4	5	5	5	5	8	9	6	5	5	4	4	4	4
−.93	4	4	4	4	4	4	5	5	5	5	8	8	6	5	5	4	4	4	4
−.92	4	4	4	4	4	5	5	5	5	5	8	8	6	5	5	4	3	4	4
−.91	4	4	4	4	4	5	5	5	5	5	8	7	6	5	4	4	3	4	4
−.90	4	4	4	4	4	5	5	5	5	5	8	7	6	5	5	4	3	4	4

Entries to the left of the dotted line represent hyperbolic orbits; those to the right of the dotted line represent elliptic orbits.

Table 7. Number of iterations required to compute x to twelve significant digits using the new method for λ near -1.

	T=0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	3.0	5.0	7.0	9.0	11.0	13.0	15.0	17.0	19.0
λ=-.99	4	5	5	5	5	5	5	5	5	5	11	14	9	6	6	5	5	4	4
-.98	4	4	5	5	5	5	5	5	5	5	10	13	8	6	6	5	5	4	4
-.97	4	5	5	5	5	5	5	5	5	5	10	12	8	6	5	5	5	4	4
-.96	4	4	5	5	5	5	5	5	5	5	9	11	7	6	5	5	4	4	4
-.95	4	4	5	5	5	5	5	5	5	5	9	10	7	6	5	5	4	4	4
-.94	4	4	5	5	5	5	5	5	5	5	9	9	7	6	5	5	4	4	4
-.93	4	4	5	5	5	5	5	5	5	5	9	9	7	6	5	5	4	4	4
-.92	4	4	5	5	5	5	5	5	5	5	9	8	6	6	5	5	4	4	4
-.91	4	5	5	5	5	5	5	5	5	5	8	8	6	6	5	5	4	4	4
-.90	4	4	5	5	5	5	5	5	5	5	8	8	6	5	5	4	4	4	4

Entries to the left of the dotted line represent hyperbolic orbits; those to the right of the dotted line represent elliptic orbits.

Table 8. Number of iterations to compute x to 8 significant digits for the minimum energy orbit when λ is near -1.

λ	T_m	Number of Iterations
-0.99	5.7207928	15
-0.98	5.4924803	12
-0.97	5.3204298	11
-0.96	5.1779966	10
-0.95	5.0547895	9
-0.94	4.9454451	9
-0.93	4.8467569	9
-0.92	4.7566233	8
-0.91	4.6735735	8
-0.90	4.5965290	8

5.2 Graphic Interpretation of the Iteration Process

The results given in the previous section show that the choice of the free parameter β described in Chapter 3 does improve the convergence properties. To see why this is so, consider the successive substitution process which involves the simultaneous solution of two equations for both Gauss' method and the new method. Figure 11(a) shows a plot of two arbitrary functions $y_1(x)$ and $y_2(x)$; the intersection of these two curves is the solution point. To find the solution by successive substitution, start by choosing an initial value x_o and finding the corresponding value $y_1(x_o)$. Next, $y_1(x_o)$ is used to obtain a new value of x by locating the point where $y_2(x) = y_1(x_o)$. The new value of x is again used to evaluate y_1 and y_1, in turn, is used to find another trial value for x. The process is repeated until an x is found for which $y_1(x) = y_2(x)$. On the graph of Fig. 11(a), this process is represented by the horizontal and vertical dotted lines. The number of iterations needed to converge to the solution point corresponds to the number of "steps" formed by these lines, each iteration consisting of one vertical and one horizontal segment. It is seen that the curvature of the functions $y_1(x)$ and $y_2(x)$ greatly influences the number of iterations. Suppose, for example, that $y_1(x)$ is a constant as shown in Fig. 11(b). In this case, the solution point would be reached in one iteration step since for any x_o, $y_1(x_o)$ equals $y_2(x)$ at the solution point.

By choosing the free parameter β such that $dy/dx = 0$ for the cubic equation (55) at the solution point, the new method seeks to flatten this curve around the solution point. Gauss' method, on the other hand, achieves a flattening of the cubic equation (15) only for small values of x near $x = 0$. To illustrate this, graphs similar to Fig. 11(a) will be presented for several pairs of λ and T for both Gauss' method and the new method. For each method, the function $y_1(x)$ is the cubic equation for y, either equation (15) or equation (55), and the function $y_2(x)$ is the defining equation for y, either equation (10) or equation (44). First consider cases of transfer angle less than 180 degrees, or positive λ. Figure 12 shows the plots for an elliptic orbit where $\lambda = +0.90$ and $T = 2$ while Fig. 13 shows plots for a hyperbolic orbit where $\lambda = +0.70$ and $T = 0.80$. These are cases where T is small, x is near 0, and Gauss' method converges rapidly. Notice the similarity in plots for each method and the near constant value of the cubic. Figure 14 shows a case of an elliptic orbit where $\lambda = +0.10$ and $T = 20$. For this large value of T, the quantity x in Gauss' method is sufficiently far from zero so that the cubic is no longer constant but slopes downward near the solution point. This causes an increase in the number of iterations for convergence. The cubic curve for the new method behaves in the opposite manner; it is sloping downward near $x = 0$ but flattens around the solution point. Plots for two cases where $\lambda = 0$ are shown for the new method in Fig. 15 Gauss' method is

singular for these cases, but the new method displays a nearly horizontal cubic curve around the solution point.

Now consider cases of transfer angle greater than 180 degrees, or negative λ. Figure 16 shows a case of an elliptic orbit where $\lambda = -0.90$ and $T = 15$. Gauss' method converges for this case but requires almost 100 iterations for eight significant figures. The reason for this is apparent from Fig. 16(b); the solution point is close to $x = 1$ where both curves are nearly vertical and gradually meet at the solution point. Clearly having the cubic equation flat near $x = 0$ is of no advantage in this case. The new method takes fewer iterations since its cubic equation is flat near the solution point. Notice that the region where the cubic of the new method has three real roots begins to appear on the graph. The method would still converge if a different x_o was chosen since the value of y on the upper branch of the cubic is chosen where possible. Two alternate iteration paths are shown in Fig. 16(c) for $x_o = 0$ and $x_o = l$. Figure 17 shows plots for a case of a hyperbolic orbit with $\lambda = -0.90$ and $T = 0.90$ where Gauss' method fails to converge. The reason for this is seen in the position of the branches of the cubic in Fig. 16(a). In this case, the iteration process oscillates in the broad region around $x = 1$ although the solution is obviously in the region $x < 0$. While the new method converges as expected, it is interesting to note that the region around the solution point where the cubic equation is flat is becoming smaller.

Figure 18 is a graph of the new method for the case of the minimum energy orbit with $\lambda = -0.95$. Here the region where the cubic is flat has condensed to a point and the curve peaks rather than flattens near the solution point. Also, the solution point has moved very close to the point of transition from one negative real root to three real roots for the cubic equation. This accounts for the increase in the number of iteration steps near the value of T for the minimum energy orbit. It is a similar peaking effect which causes problems for the Newton-Rapheson approach. Lancaster and Blanchard[19] give a plot of T versus the chosen iterated variable for λ near -1 in which the curves are approaching a cusp in the neighborhood of the minimum energy orbit. The first derivative used in the Newton-Rapheson iteration process ceases to exist for $\lambda = -1$ at the minimum energy orbit.

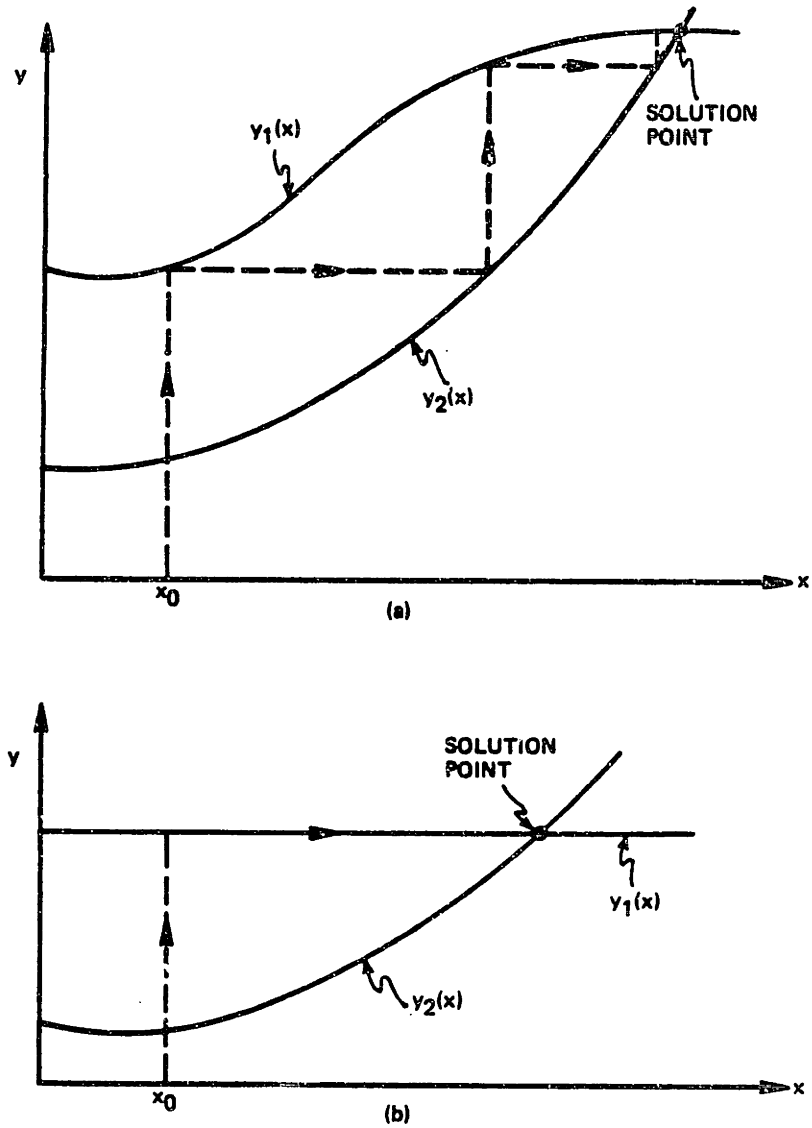
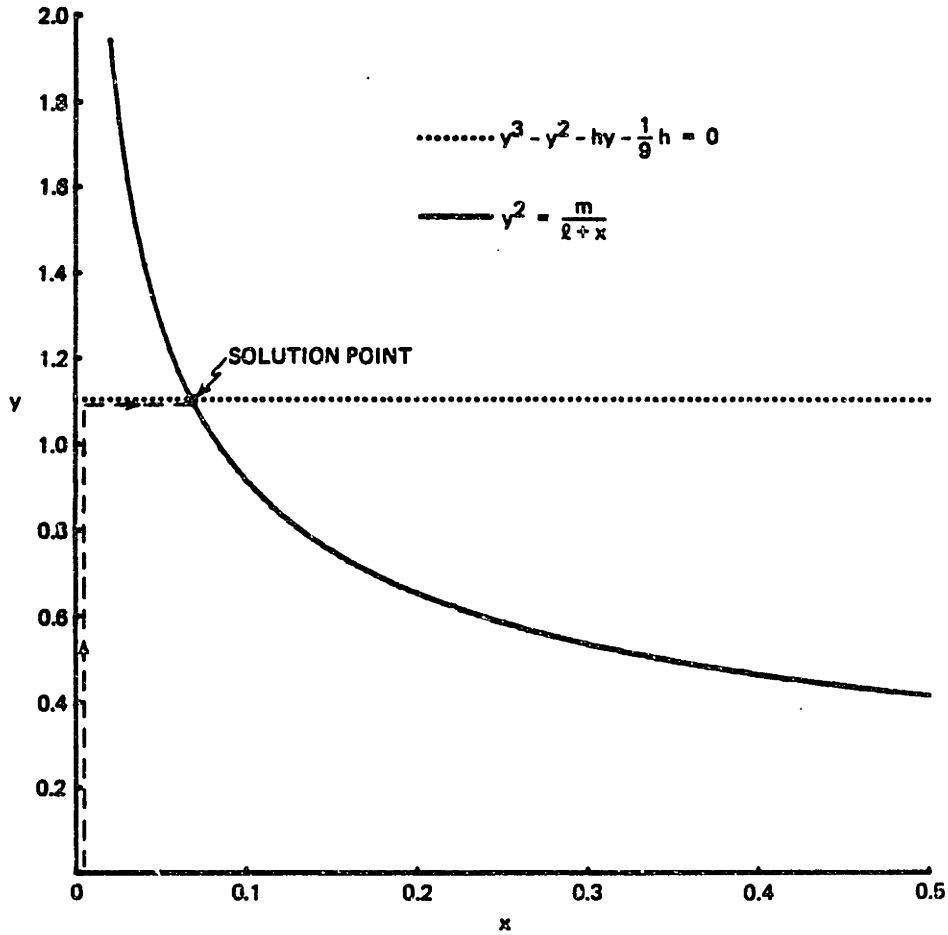
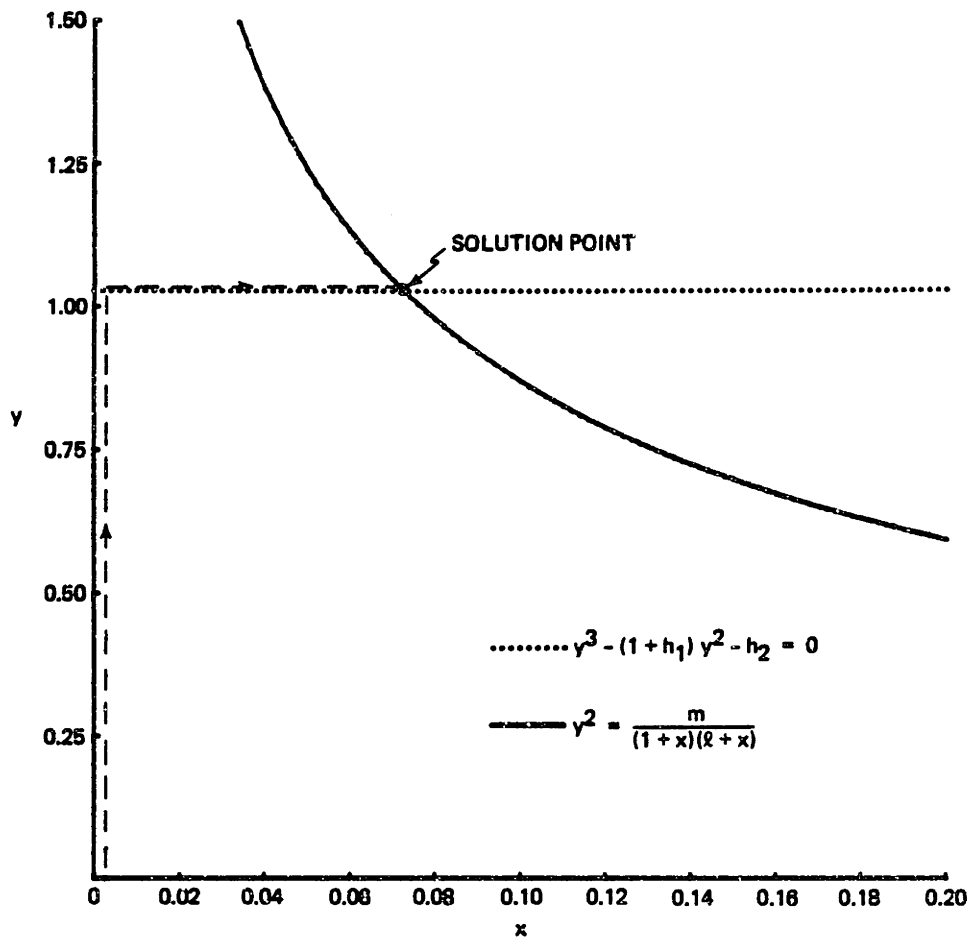


Figure 11. Graphic Interpretation of the Iteration Process



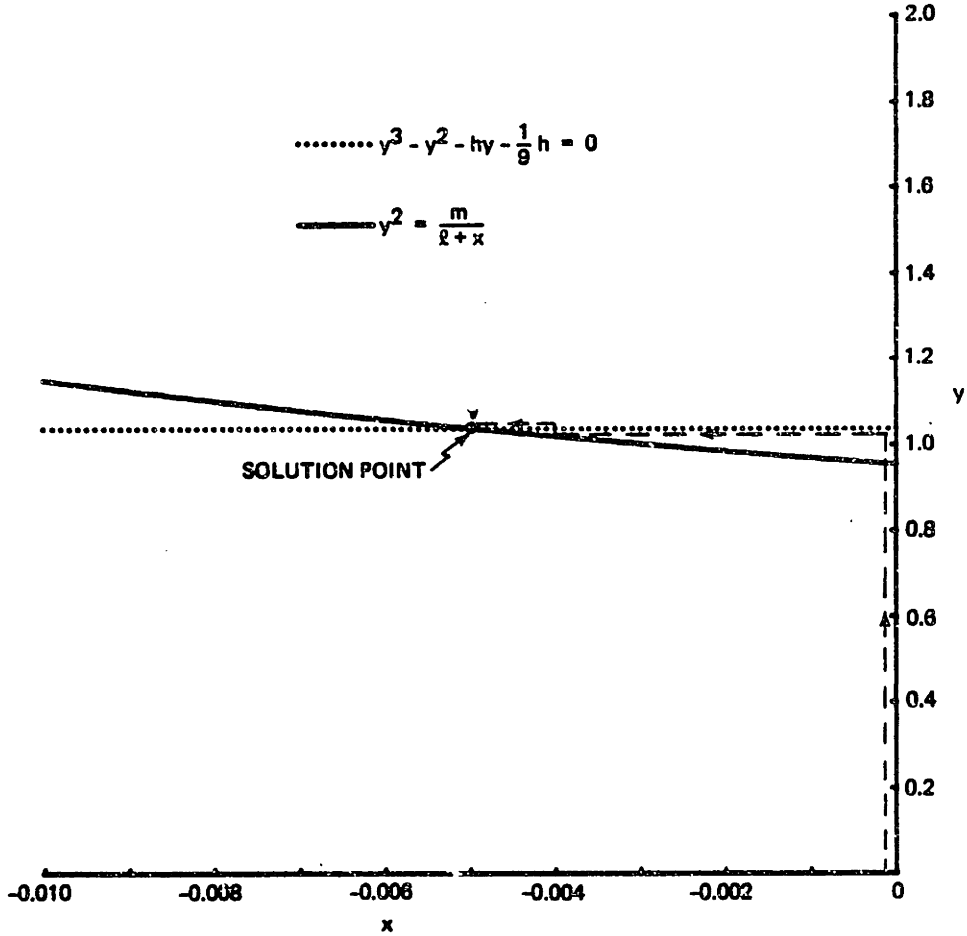
(a) Gauss' method

Figure 12. Iteration Paths for the Elliptic Orbit with $\lambda = +0.9$ and $T = 2$



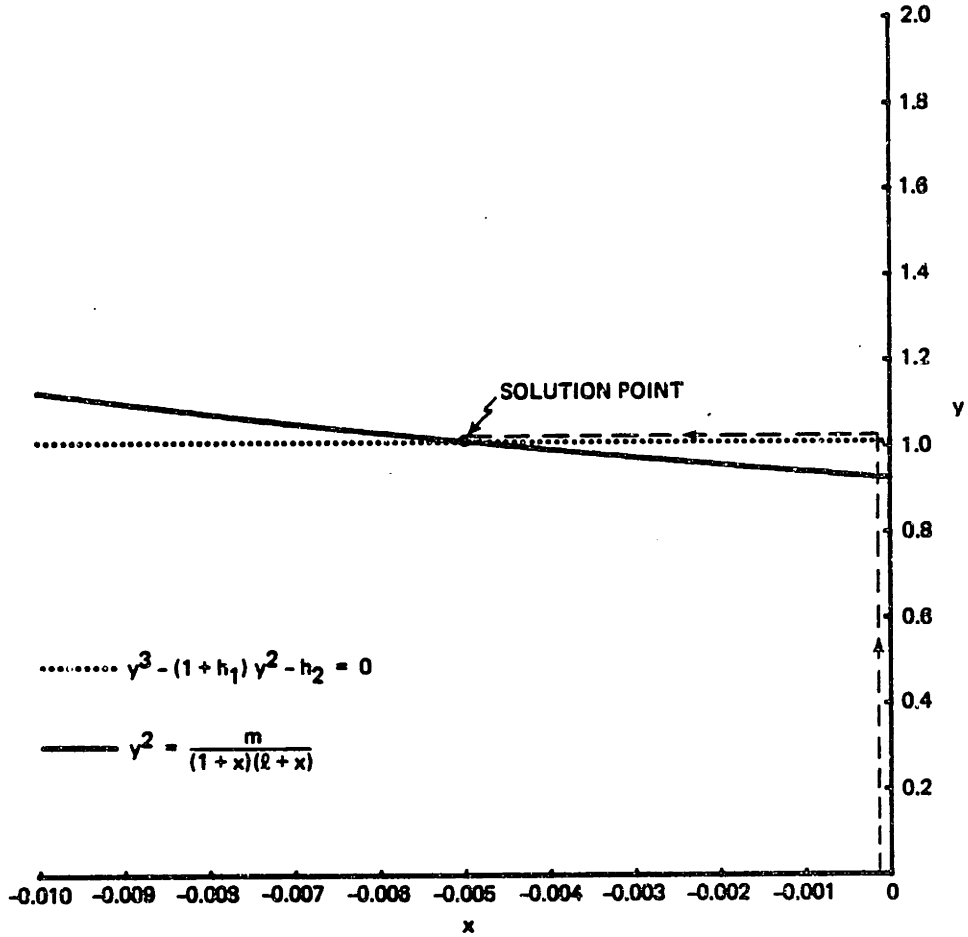
(b) New method

Figure 12. Iteration Paths for the Elliptic Orbit with $\lambda = +0.9$ and $T = 2$ (Cont'd)



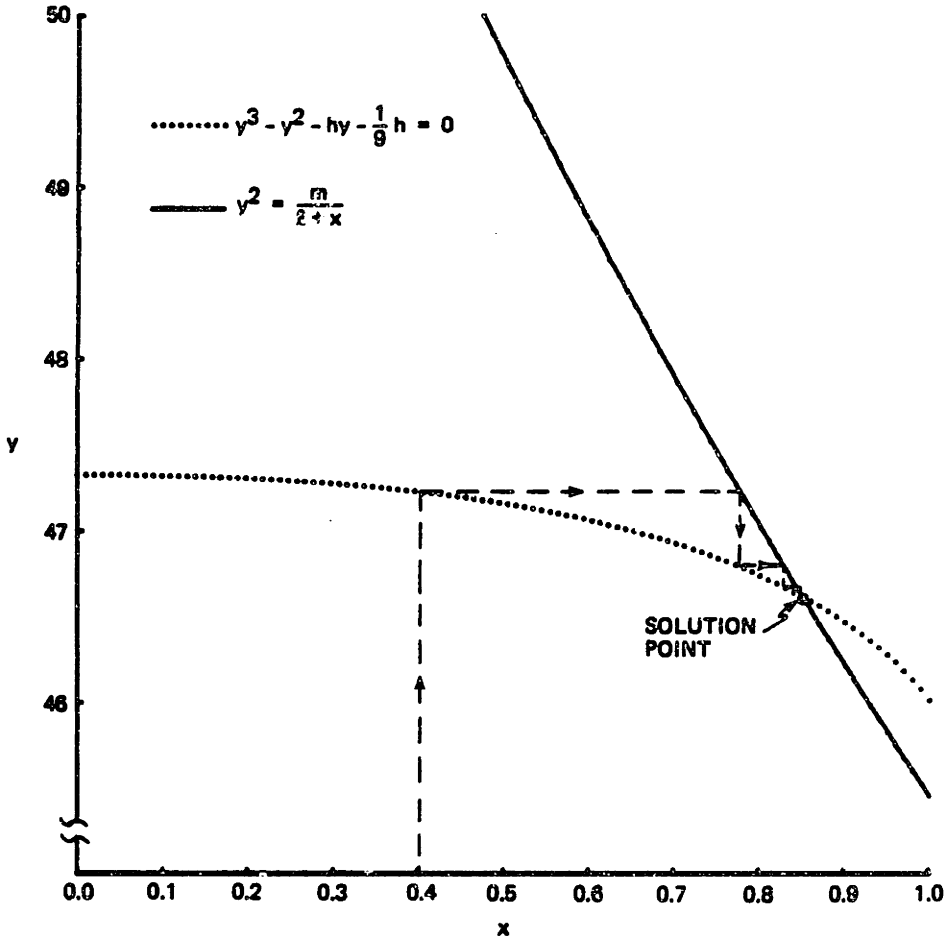
(a) Gauss' method

Figure 13. Iteration Paths for the Hyperbolic Orbit with $\lambda = +0.7$ and $T = 0.80$



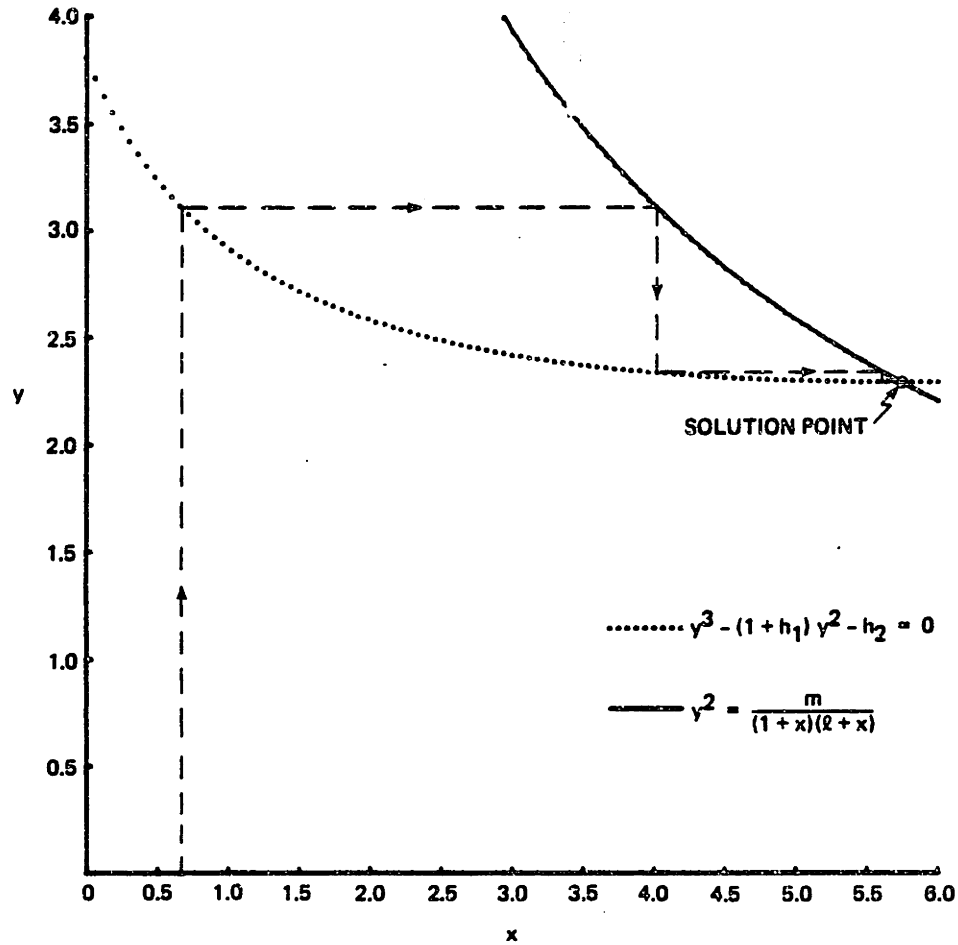
(b) New method

Figure 13. Iteration Paths for the Hyperbolic Orbit with $\lambda = +0.7$ and $T = 0.80$ (Cont'd)



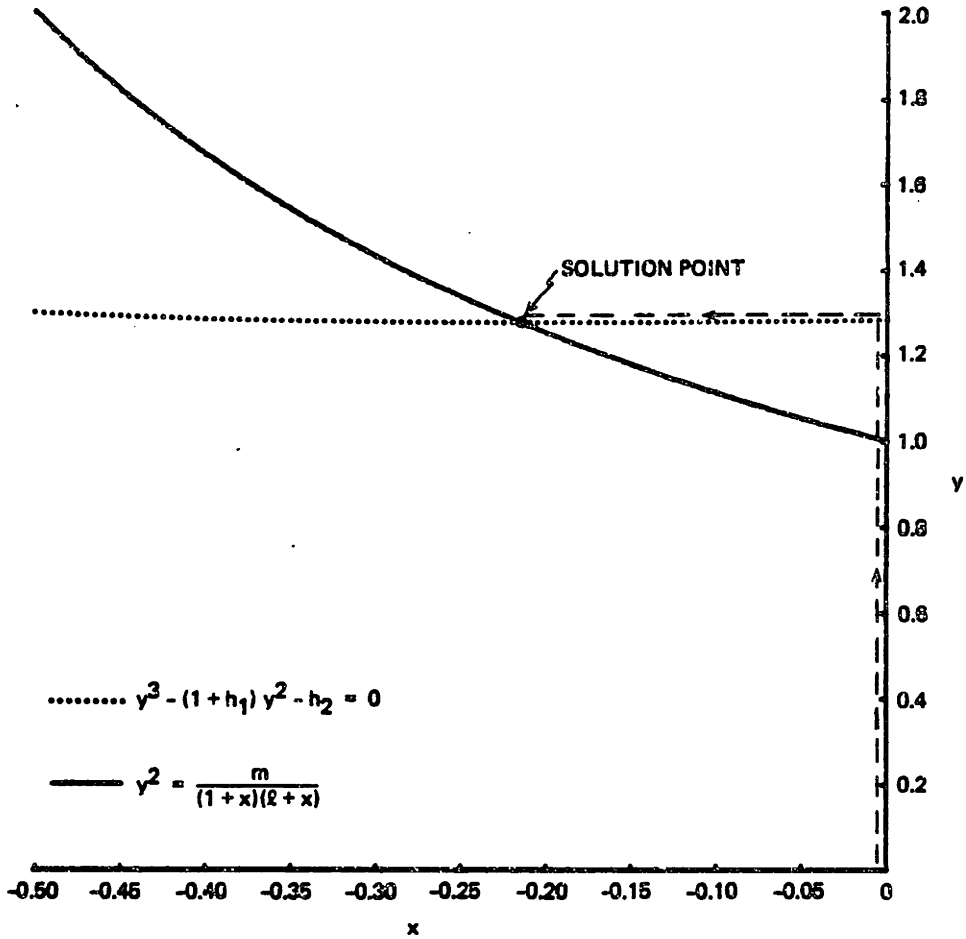
(a) Gauss' method

Figure 14. Iteration Paths for the Elliptic Orbit with $\lambda = +0.1$ and $T = 20$



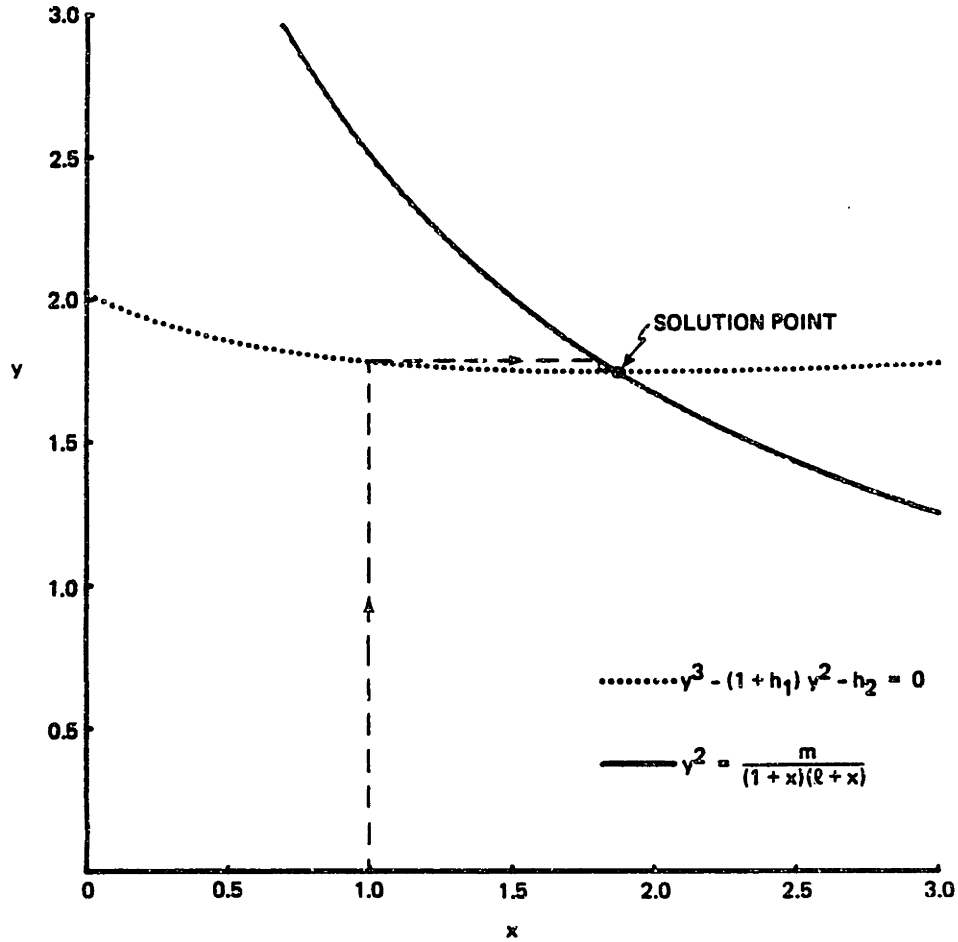
(b) New method

Figure 14. Iteration Paths for the Elliptic Orbit with $\lambda = +0.1$ and $T = 20$ (Cont'd)



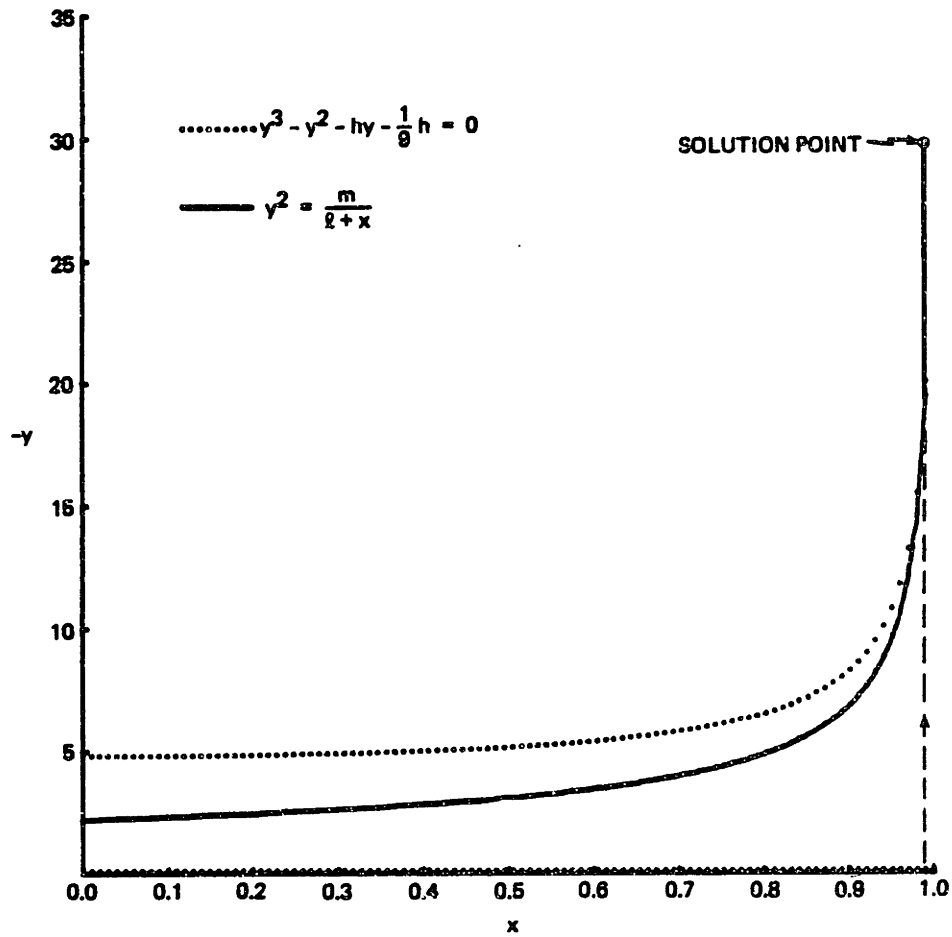
(a) Hyperbolic orbit with T = 1.0

Figure 15. Iteration Paths for Orbits with $\lambda = 0$ Using the New Method



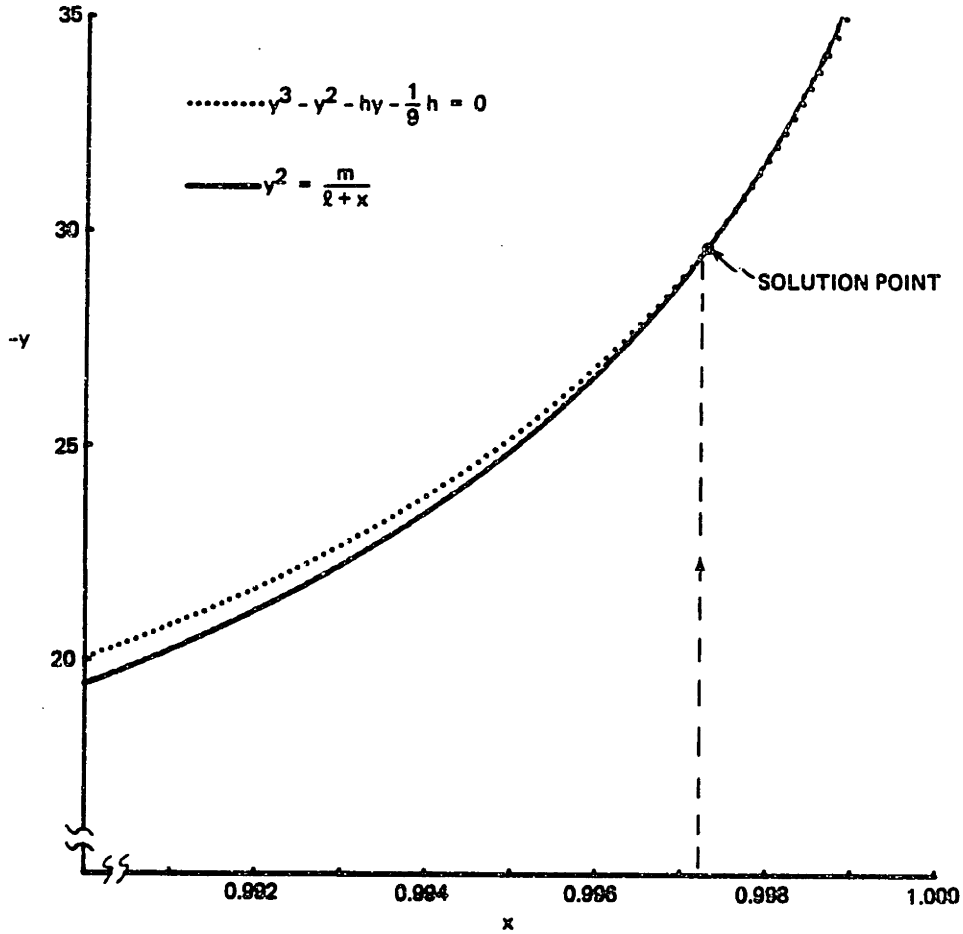
(b) Elliptic orbit with T = 5.0

Figure 15. Iteration Paths for Orbits with $\lambda = 0$ Using the New Method (Cont'd)



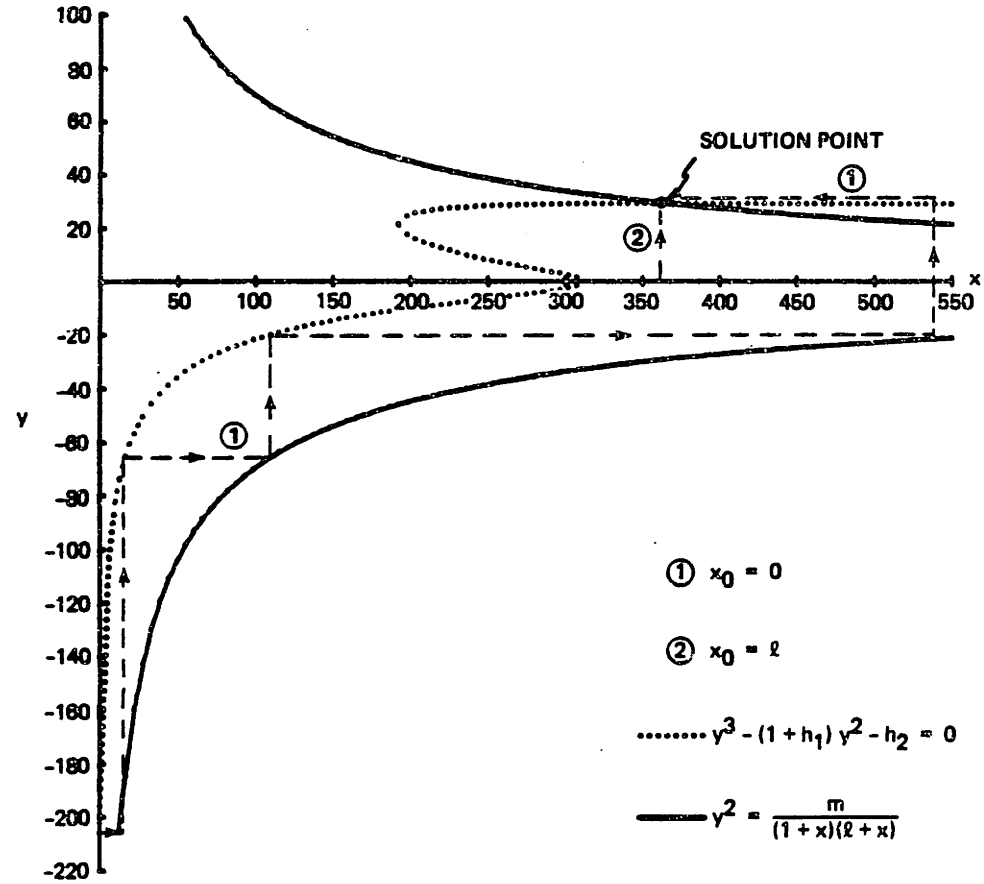
(a) Curves of Gauss' method for $0 < x < 1$

Figure 16. Iteration Paths for the Elliptic Orbit with $\lambda = -0.9$ and $T = 15$



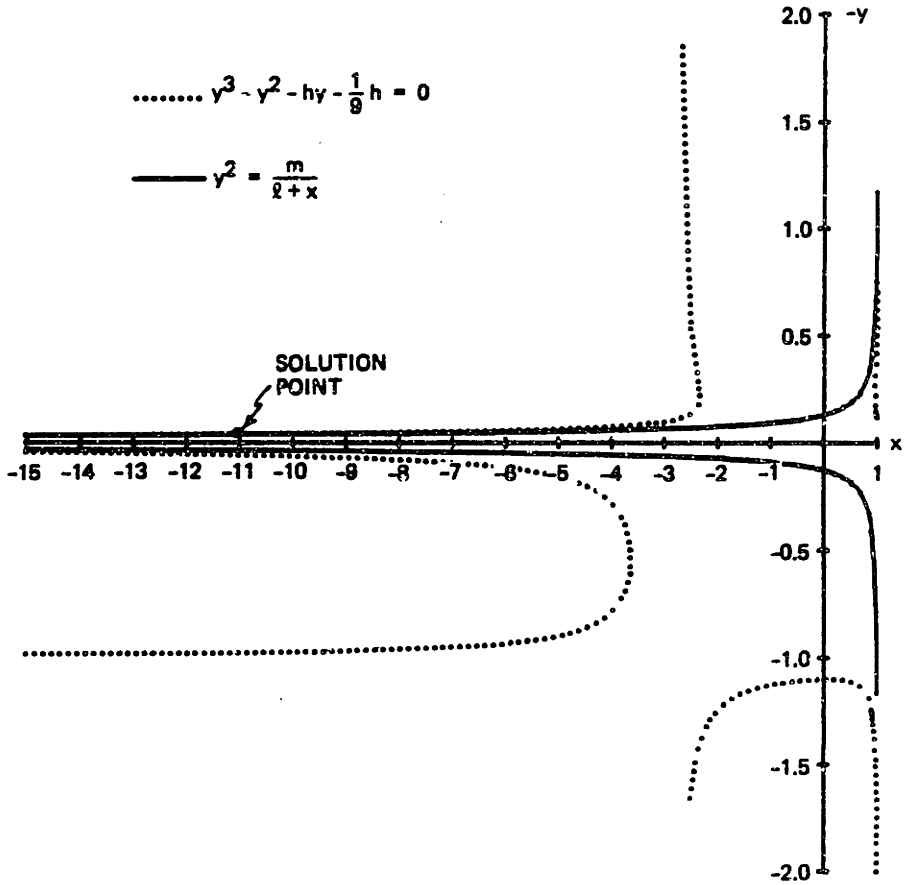
(b) Curves of Gauss' method for $0.9 < x < 1.0$

Figure 16. Iteration Paths for the Elliptic Orbit with $\lambda = -0.9$ and $T = 15$ (Cont'd)



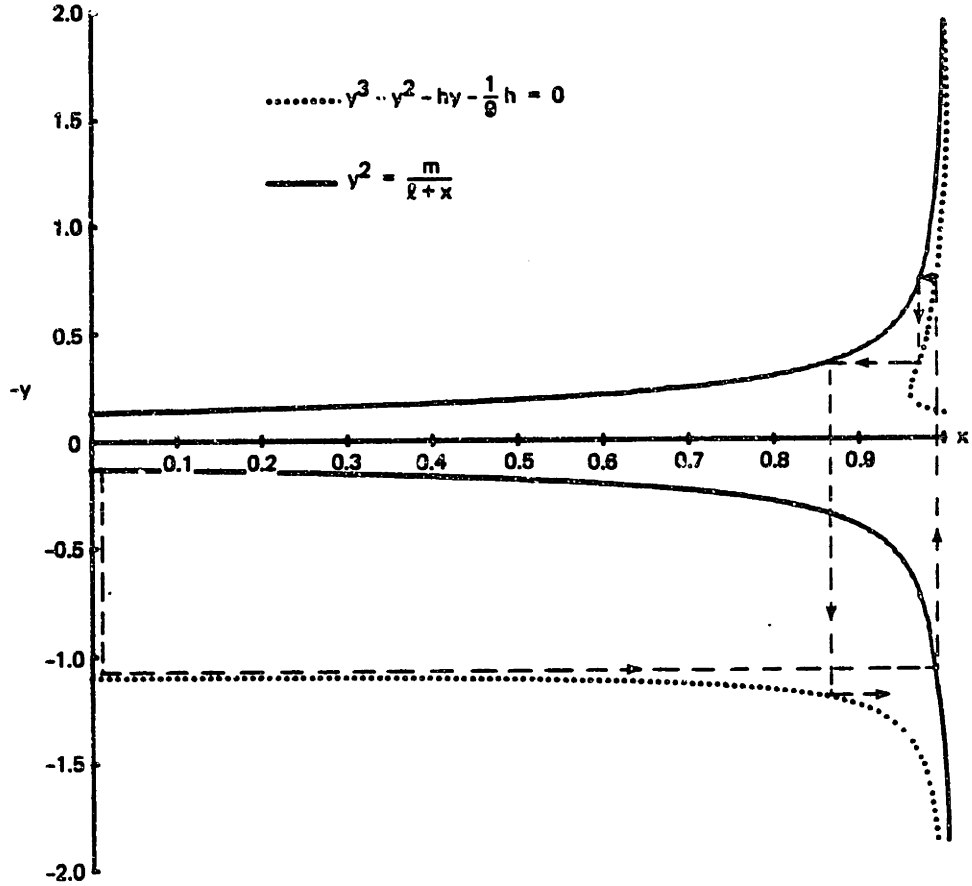
(c) Curves for the new method showing two iteration paths, $x_0 = 0$ and $x_0 = \ell$

Figure 16. Iteration Paths for the Elliptic Orbit with $\lambda = -0.9$ and $T = 15$ (Cont'd)



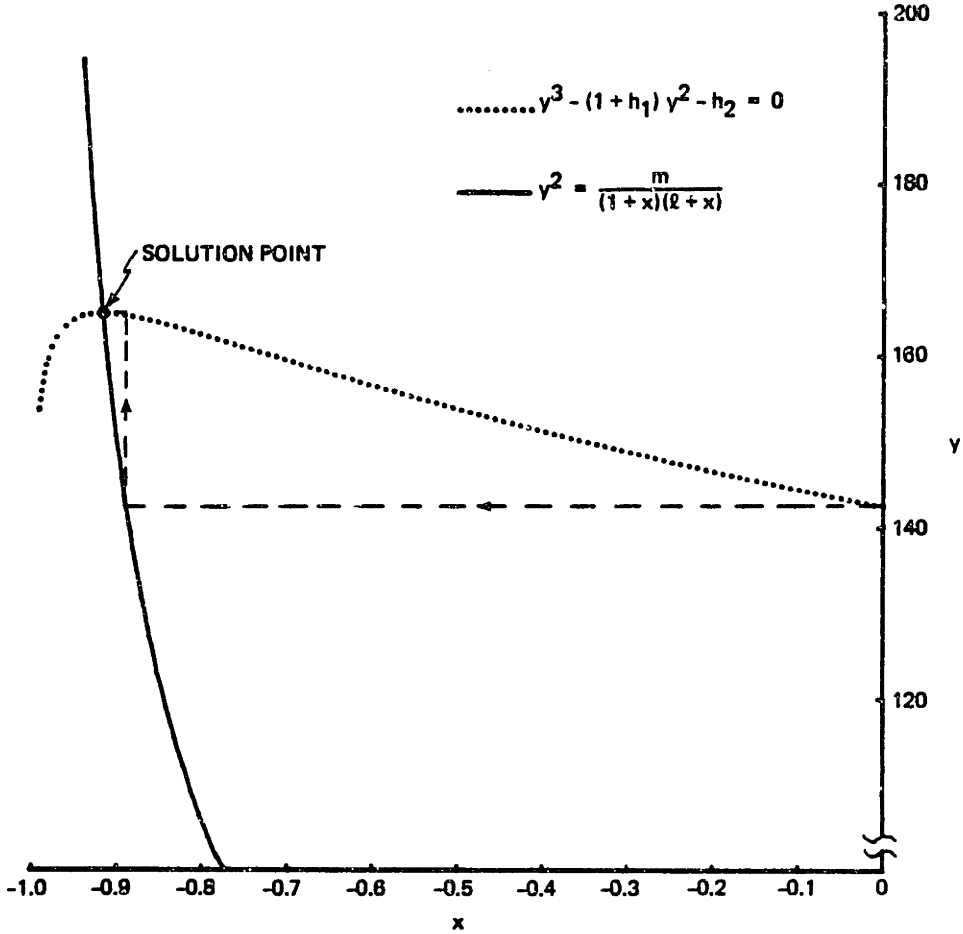
(a) Curves for Gauss' method showing the solution point

Figure 17. Iteration Paths for the Hyperbolic Orbit with $\lambda = -0.9$ and $T = 0.90$



(b) Iteration path for Gauss' method

Figure 17. Iteration Paths for the Hyperbolic Orbit with $\lambda = -0.9$ and $T = 0.90$ (Cont'd)



(c) New method

Figure 17. Iteration Paths for the Hyperbolic Orbit with $\lambda = -0.9$ and $T = 0.90$ (Cont'd)

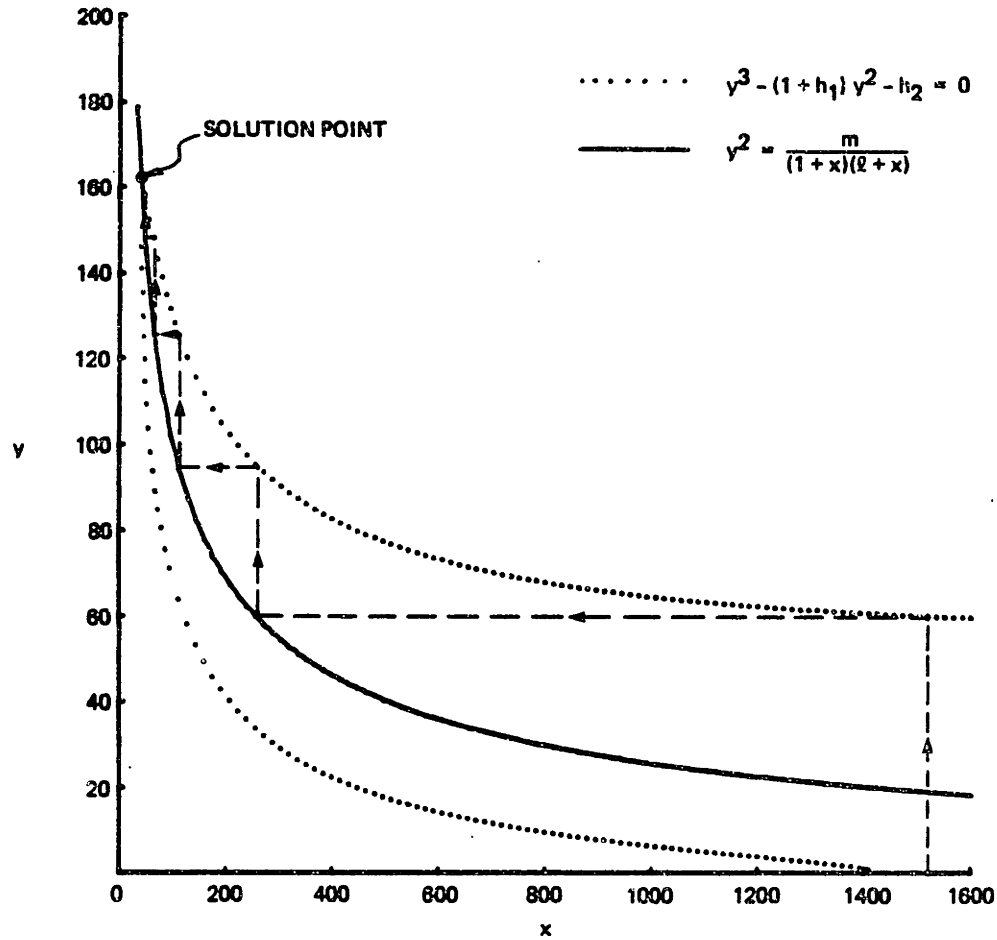


Figure 18. Iteration Path for the Minimum Energy Orbit with $\lambda = -0.95$ Using the New Method

Chapter 6

Conclusions

THE NEW METHOD of solving Lambert's problem presented in the previous chapters closely parallels Gauss' elegant method, but is superior to it. This method is designed to converge rapidly for any given geometry and time of flight. It retains all the original advantages of Gauss' method and is applicable to a wide range of problems of practical interest. The convergence properties of the iteration process are nearly uniform; the amount of computation needed to compute a solution to any specified accuracy is about the same for all cases. A major advantage is that only a small increase in the number of iterations is required to increase the precision of the solution.

The only restriction in using this method is that it is singular for the case of a 360-degree transfer angle. The convergence properties are not greatly affected by this singularity even in cases which are close to a 360-degree transfer. It should be noted that the singularity for 360-degree transfer angles occurs for most Newton-Rapheson schemes. Also, the new method can be used without modification in cases near the 360-degree transfer angle minimum energy orbit where a Newton-Rapheson scheme must be abandoned.

In most guidance applications, this method appears to be an attractive alternative to present Newton-Rapheson iteration schemes for solving Lambert's problem. There is no need to find specific starting values for different input parameters. The best initial guess was found to be the parabolic orbit for a parabola or a hyperbola and a circular orbit for an ellipse. If Lambert's problem is to be solved at sequential intervals of time, additional computation time could be saved by using the solution from the previous time step as the starting guess for the present time. Further testing is recommended to better compare this new method to currently implemented solutions of Lambert's problem.

References

[1]Volk O., "Johann Heinrich Lambert and the Determination of Orbits for Planets and Comets," *Celestial Mechanics*, Vol. 21, 1980, pp. 237–250.

[2]Bate, Roger R., Mueller, Donald D., and White, Jerry E., *Fundamental Astrodynamics*, Dover Publications, Inc., New York, 1971, pp. 227–228.

[3]Bell, Eric T., *Men of Mathematics*, Dover Publications, Inc., New York, 1937, p. 239.

[4]Battin, Richard H., *Astronautical Guidance*, McGraw-Hill Book Co., Inc., New York, 1964.

[5]Battin, Richard H., "A New Solution for Lambert's Problem," XIXth International Astronautical Congress, Vol. 2 Astrodynamics and Astrionics, Pergamon Press, Ltd., New York, 1970.

[6]Battin, Richard H., "Lambert's Problem Revisited," *AIAA Journal*, Vol. 15, No. 5, May 1977, pp. 707–713.

[7]Gauss, Carl Fredrich, *Theory of Motion of the Heavenly Bodies Moving about the Sun in Conic Sections*, a translation of *Theoria Motus* by C.H. Davis, Dover Publications, Inc., New York, 1963, pp. 114–145.

[8]Deprit, Andre, *Fundamentals of Astrodynamics (Part I)*, Mathematical Note No. 556, Mathematics Research Laboratory, Boeing Scientific Research Laboratories, April 1968, p. 187.

[9]Gauss, Carl Fredrich, *op. cit.*, p. 140.

[10]Battin, Richard H., *Astronautical Guidance*, p. 21.

[11]Battin, Richard H., Course lecture for 16.47 - Astronautical Guidance, M.I.T., April 12, 1982.

[12]Battin, Richard H., Fill, Thomas J., and Shepperd, Stanley W., "A New Transformation Invariant in the Orbital Boundary-Value Problem," *Journal of Guidance and Control*, Vol. 1, Jan.-Feb. 1978, pp. 50–55.

[13]Oberhettinger, F., " Hypergeometric Functions," *Handbook of Mathematical Functions* (ed. by M. Abramowitz and I. A. Stegun), Dover Publications, Inc., New York, 1965, pp. 556–566.

[14]Oberhettinger, F., *ibid.*

[15]Gautschi, W., "Computational Aspects of Three-Term Recurrence Relations," *SIAM Review*, Vol. 9, Jan. 1967, pp. 29–30.

[16]Newmark, S., *Solution of Cubic and Quartic Equations*, Pergamon Press, Ltd., New York, 1965, pp. 5–11, 31–34.

[17]Chrystal, G., *Textbook of Algebra, Part 2*, Dover Publications, Inc., New York, 1961, pp. 523–524.

[18]Lancaster, E. R., Blanchard, E. C., and Devaney, R. A., "A Note on Lambert's Theorem," *Journal of Spacecraft and Rockets*, 3, No. 9, Sept. 1966, pp. 1436–1438.

[19]Lancaster, E. R., and Blanchard, E. C., "A Unified Form of Lambert's Theorem," NASA TN D-5368, Sept. 1969.