

Advances in Industrial Control

Springer

London

Berlin

Heidelberg

New York

Barcelona

Budapest

Hong Kong

Milan

Paris

Santa Clara

Singapore

Tokyo

Other titles published in this Series:

Microcomputer-Based Adaptive Control Applied to Thyristor-Driven D-C Motors
Ulrich Keuchel and Richard M. Stephan

Expert Aided Control System Design
Colin Tebbutt

Modeling and Advanced Control for Process Industries, Applications to Paper Making Processes

Ming Rao, Qijun Xia and Yiquan Ying

Modelling and Simulation of Power Generation Plants

A.W. Ordys, A.W. Pike, M.A. Johnson, R.M. Katebi and M.J. Grimble

Model Predictive Control in the Process Industry

E.F. Camacho and C. Bordons

H_∞ Aerospace Control Design: A VSTOL Flight Application

R.A. Hyde

Neural Network Engineering in Dynamic Control Systems

Edited by Kenneth Hunt, George Irwin and Kevin Warwick

Neuro-Control and its Applications

Sigeru Omatsu, Marzuki Khalid and Rubiyah Yusof

Energy Efficient Train Control

P.G. Howlett and P.J. Pudney

Hierarchical Power Systems Control: Its Value in a Changing Industry

Marija D. Ilic and Shell Liu

System Identification and Robust Control

Steen Tøffner-Clausen

Genetic Algorithms for Control and Signal Processing

K.F. Man, K.S. Tang, S. Kwong and W.A. Halang

Advanced Control of Solar Plants

E.F. Camacho, M. Berenguel and F.R. Rubio

Control of Modern Integrated Power Systems

E. Mariani and S.S. Murthy

Advanced Load Dispatch for Power Systems: Principles, Practices and Economies

E. Mariani and S.S. Murthy

Supervision and Control for Industrial Processes

Björn Sohlberg

Modelling and Simulation of Human Behaviour in System Control

Pietro Carlo Cacciabue

Modelling and Identification in Robotics

Krzysztof Kozłowski

Maxwell Noton

Spacecraft Navigation and Guidance

With 34 Figures

Software associated with this text is available free to the reader by means of the Internet server of the publisher. See Preface for details.



Springer

Dr Maxwell Noton
Independent Consultant, formerly of British Aerospace - Space Systems and NASA JPL.

British Library Cataloguing in Publication Data

Noton, Maxwell

Spacecraft navigation and guidance. - (Advances in
industrial control)

1.Navigation (Astronautics) 2.Orbital mechanics

I.Title

629.4'742

Library of Congress Cataloging-in-Publication Data

Noton, Maxwell, 1928-

Spacecraft navigation and guidance / Maxwell Noton.

p. cm. -- (Advances in industrial control)

ISBN-13: 978-1-4471-1585-4 e-ISBN-13: 978-1-4471-1583-0

DOI: 10.1007/978-1-4471-1583-0

1. Navigation (Astronautics) 2. Space vehicles--Guidance systems.

I. Title. II. Series.

TL1065.N68 1998

97-53165

629.4'742--dc21

CIP

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

© Springer-Verlag London Limited 1998

Softcover reprint of the hardcover 1st edition 1998

The use of registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant laws and regulations and therefore free for general use.

The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

Typesetting: Camera ready by author

69/3830-543210 Printed on acid-free paper

Advances in Industrial Control

Series Editors

**Professor Michael J. Grimble, Professor of Industrial Systems and Director
Dr. Michael A. Johnson, Reader in Control Systems and Deputy Director**

**Industrial Control Centre
Department of Electronic and Electrical Engineering
University of Strathclyde
Graham Hills Building
50 George Street
Glasgow G1 1QE
United Kingdom**

Series Advisory Board

**Professor Dr-Ing J. Ackermann
DLR Institut für Robotik und Systemdynamik
Postfach 1116
D82230 Weßling
Germany**

**Professor I.D. Landau
Laboratoire d'Automatique de Grenoble
ENSIEG, BP 46
38402 Saint Martin d'Hères
France**

**Dr D.C. McFarlane
Department of Engineering
University of Cambridge
Cambridge CB2 1QJ
United Kingdom**

**Professor B. Wittenmark
Department of Automatic Control
Lund Institute of Technology
PO Box 118
S-221 00 Lund
Sweden**

**Professor D.W. Clarke
Department of Engineering Science
University of Oxford
Parks Road
Oxford OX1 3PJ
United Kingdom**

Professor Dr -Ing M. Thoma
Westermannweg 7
D-30419 Hannover
Germany

Professor H. Kimura
Department of Mathematical Engineering and Information Physics
Faculty of Engineering
The University of Tokyo
7-3-1 Hongo
Bunkyo Ku
Tokyo 113
Japan

Professor A.J. Laub
College of Engineering - Dean's Office
University of California
One Shields Avenue
Davis
California 95616-5294
United States of America

Professor J.B. Moore
Department of Systems Engineering
The Australian National University
Research School of Physical Sciences
GPO Box 4
Canberra
ACT 2601
Australia

Dr M.K. Masten
Texas Instruments
2309 Northcrest
Plano
TX 75075
United States of America

Professor Ton Backx
AspenTech Europe B.V.
De Waal 32
NL-5684 PH Best
The Netherlands

SERIES EDITORS' FOREWORD

The series *Advances in Industrial Control* aims to report and encourage technology transfer in control engineering. The rapid development of control technology impacts all areas of the control discipline. New theory, new controllers, actuators, sensors, new industrial processes, computer methods, new applications, new philosophies,..., new challenges. Much of this development work resides in industrial reports, feasibility study papers and the reports of advanced collaborative projects. The series offers an opportunity for researchers to present an extended exposition of such new work in all aspects of industrial control for wider and rapid dissemination.

One of the defining technologies of the 20th Century is that of Space Flight. Control Systems techniques contributed to this engineering development in many ways but the most cited contribution is that of Kalman filtering. In Maxwell Noton's monograph the applications context of Kalman filtering is carefully presented in the wider science of Spacecraft Navigation and Guidance.

This very readable monograph describes and mathematically models the stages of a mission through the Launch phase, into Mid-course and Gravity Assist Manoeuvres and finally to Atmospheric Entry. The mathematical descriptions are accompanied by reference to many space missions of the last thirty years or so. A novel feature is the availability of C++ programs and appropriate data to be able to repeat computational examples given. This facility is likely to be of particular interest to graduate and postgraduate students following courses in these spacecraft navigation topics. The chapter on Orbit Determination demonstrates the progression of numerical techniques as the limitations of Least Squares parameter estimation, led to the Square Root Information Filter, and onto Sequential Parameter Estimation and the Kalman filter.

The clarity of the monograph with its careful aggregation of mathematical models for the various mission stages makes this monograph a good source for the basis of Spacecraft Navigation. We hope that it finds a wide readership in the astronautic and control engineering community.

M.J. Grimble and M.A. Johnson
Industrial Control Centre
Glasgow, Scotland, UK

PREFACE

The title of this monograph might simply have been "Spacecraft Navigation" but *Navigation and Guidance* has been adopted in line with current usage to underline that the text is concerned both with navigation (determining current position and velocity or equivalent variables) and guidance (correcting the present trajectory). Having clarified these terms, it must be admitted that no attempt has been made to restrict the scope of the text unduly, with the result that some of the included material could be categorised as *mission analysis*.

In one respect the treatment might be regarded as limited because there is little attention to onboard autonomous systems of navigation and guidance. This may seem surprising to some readers but the fact is that most spacecraft navigation to date has been by means of ground-based data processing, computation , and radio commands to the spacecraft. This has been the simpler, cheaper and reliable option. Exceptions to this arise when communication times between spacecraft and ground cannot be tolerated (eg. landing on a planetary body) or when it is necessary to enhance human safety such as a manned landing on the Moon or return to Earth.

The text differs to some extent from earlier treatments because *computer methods* have been recognised , emphasized and their importance illustrated. Such technology was not of course available to famous mathematicians of bygone era, but the author did not feel constrained to devise contrived analytical examples and questions to test the manipulative ability of students. Instead it has been assumed that the reader has available to him a typical modern personal computer. If the text were used in conjunction with a graduate course, then it is assumed that student assignments would take the form of computer-oriented exercises of varying scope.

With this in mind, the author has employed *software written specifically to illustrate the developments in each chapter*. The suites of computer programs are modular and written in C++, recognising the importance of that more powerful language despite the amount of similar FORTRAN programs that have developed over the years by engineers. All this C++ software has been deposited in the public server of the publisher. It is available free to the reader, but without support, simply by adapting the FTP mode in any gateway to the Internet and downloading the file NOTON.ZIP (256,201 bytes) from the

publisher's site [ftp.springer.de](ftp://ftp.springer.de) in the directory /pub/technik/noton. After deflating the single big file with an unzip utility, all the data files and source programs will be available as listed at the end of each chapter.

CONTENTS

Preface	IX
1. Orbital Mechanics	1
1.1 Introduction	1
1.2 The Two-Body Problem	2
1.3 Orbital Elements and Reference Axes	4
1.3.1 Position and Velocity from the Orbital Elements	6
1.3.2 Orbital Elements from the Position and Velocity	7
1.3.3 Numerical Examples	8
1.4 Time in Orbit	9
1.4.1 Solution for Ellipse	9
1.4.2 Solution for Hyperbola	11
1.4.3 Computations starting from Time in Orbit	11
1.5 Lambert's Time of Flight Theorem	12
1.6 Orbital Manoeuvres	16
1.6.1 Coplanar Transfer Manoeuvres	16
1.6.2 Injection into an Interplanetary Orbit	17
1.6.3 Plane Change to inject into a Geostationary Orbit	19
1.7 Patched Conics	20
1.7.1 Spheres of Influence	20
1.7.2 Time in Julian Days	21
1.7.3 A Solar Polar Mission	22
1.8 Numerical Integration of Orbits	25
1.8.1 Cowell's Method	25
1.8.2 Encke's Method	26
1.9 Software	27
2. The Launch Phase	29
2.1 Introduction	29
2.2 Propulsion	30
2.2.1 Basics	30
2.2.2 Staging	31
2.3 Launch Systems	33
2.4 Trajectory Dynamics	35

2.4.1	Equations of Motion	35
2.4.2	Gravitational Forces	38
2.4.3	Rocket Thrust	39
2.4.4	Aerodynamic Forces	39
2.4.5	Final Orbital Elements	41
2.5	Computation of the Nominal Ascent Trajectory	41
2.6	A Computed Example: Ariane 4	44
2.7	Optimization of the Ascent	47
2.7.1	Solution for the Flat Earth	47
2.7.2	Improvement of the Nominal Trajectory	49
2.8	In-flight Guidance	51
2.9	The Guidance and Navigation System	52
2.10	Software	54
3.	Earth Satellites: Perturbations and Manoeuvres	55
3.1	Introduction	55
3.2	Variation of the Orbital Elements.....	55
3.2.1	Rates of Change of the Elements	55
3.2.2	Use of Radial and Transverse Components.....	58
3.2.3	Use of Tangential and Normal Components	59
3.2.4	Summary of Equations in Radial-Transverse Axes	60
3.2.5	Summary of Equations in Tangential-Normal Axes	60
3.2.6	Equations Applicable at Zero Eccentricity	60
3.2.7	The Singularity at Zero Inclination	61
3.3	Gravitational Field of the Earth	62
3.3.1	Expansion of Spherical Harmonics	62
3.3.2	Evaluation for an Oblate Earth	63
3.3.3	Analytical Results for an Oblate Earth	64
3.4	Perturbations from the Sun and the Moon	66
3.5	Effect of Atmospheric Drag	67
3.6	Illustrative Computations of Perturbations	68
3.7	Correcting Low Earth Orbits	68
3.8	Geostationary Orbits	69
3.8.1	Introduction	69
3.8.2	North-South Perturbations due to the Moon and Sun	70
3.8.3	East-West Perturbations due to Triaxial Effects	71
3.8.4	East-West Perturbations due to Solar Pressure	73
3.9	Rendezvous Guidance	75
3.9.1	Introduction	75
3.9.2	Guidance Strategy	76
3.9.3	Examples by Simulation	77
3.9.4	Other references to Rendezvous and Docking	79
3.10	Software	80

4. Orbit Determination	81
4.1 Introduction	81
4.2 First Estimates of Orbits	83
4.2.1 Transformation to Observables	83
4.2.2 Planetary Orbits	85
4.2.3 Computer Examples of Planetary Orbits	87
4.2.4 Spacecraft Orbits	88
4.2.5 Computer Example of Earth Satellite	90
4.3 Refinement of Orbits	90
4.3.1 Batch Least-Squares Fitting	90
4.3.2 Generation of Partial Derivatives	92
4.3.3 The Square Root Information Filter	93
4.3.4 Bias Terms and the Consider Matrix	96
4.4 Sequential Estimation	97
4.4.1 Non-constant Parameters	97
4.4.2 Time Update with Sequential SRIF	98
4.4.3 Measurement Update with sequential SRIF	100
4.4.4 Sequential Estimation with Covariance Filters	100
4.4.5 SRIF Example with an Earth Satellite Orbit	101
4.4.6 SRIF Example for a Spacecraft Approaching a Comet .	104
4.5 Software	108
5. Midcourse and Gravity Assist Manoeuvres	109
5.1 Introduction	109
5.2 Midcourse Manoeuvres	110
5.2.1 Nominal Trajectory to Jupiter	110
5.2.2 Errors at Heliocentric Injection	111
5.2.3 Fixed Time Guidance	112
5.2.4 Variable Time Guidance	113
5.2.5 Parametric Results	115
5.3 Gravity Assist Manoeuvres	116
5.3.1 Introduction	116
5.3.2 Gravity Assist in Two Dimensions	118
5.3.3 Gravity Assist in Three Dimensions	119
5.3.4 An Example : Comet Nucleus Sample Return	121
5.4 Software	123
6. Low Thrust Missions	125
6.1 Introduction	125
6.2 Electric Propulsion	125
6.3 Optimization of Low Thrust Trajectories	126
6.3.1 Optimization in Two Dimensions	126
6.3.2 Parameterized Optimization in Three Dimensions .	129
6.3.3 Numerical Example	132
6.3.4 Controllability	133

XIV Table of Contents

6.4	Guidance on Low Thrust Trajectories	134
6.4.1	Introduction	134
6.4.2	Analysis	135
6.4.3	The Linear Quadratic Control Problem	136
6.4.4	Numerical Example of Feedback Guidance	137
6.5	Software	139
7.	Atmospheric Entry	141
7.1	Introduction	141
7.2	Flight Equations	142
7.2.1	Equations of Motion	142
7.2.2	Aerodynamic Lift and Drag	143
7.2.3	Heating	144
7.3	Entry Corridors	145
7.3.1	Parameters	145
7.3.2	Condition for Level Flight	148
7.3.3	The Limit of Total Acceleration	151
7.3.4	The Limit of Rate of Heating	151
7.4	Ballistic Entry with No Lift	152
7.5	Guided Entry using Lift	155
7.5.1	Lift and the Guidance System	155
7.5.2	The Nominal Trajectory	156
7.5.3	Feedback Guidance in the Upper Atmosphere	158
7.5.4	Feedback Guidance in the Lower Atmosphere	161
7.6	Software	164
A.	Appendices	169
A.1	Perturbations from the Sun and Moon in Geostationary Orbit ..	169
A.2	Perturbations due to Solar Pressure in Geostationary Orbit ..	170
A.3	Angular Information from Doppler Tracking	172
A.4	Ephemeris of the Sun	172
A.5	Sequential UD Covariance Filter	173
A.6	Global Positioning System	175

CHAPTER 1

ORBITAL MECHANICS

1.1 Introduction

The results of orbital mechanics required in this text can be obtained from classical mechanics and only in very exceptional circumstances might a relativistic correction be necessary. Newton's laws of motion may be combined with the universal law of gravitation to derive the motion of planets and other bodies in the solar system, and man-made satellites (spacecraft) relative to such bodies. The law of gravitation is stated for particles but it is extended by integration to bodies of finite size. However, the gravitational field outside a spherically symmetrical body is as if the mass were concentrated at the centre. For non-spherical bodies the effects of an irregular shape diminish with distance and can be ignored in most but not all cases, eg. the orbit of a low altitude satellite is perturbed by the oblate shape of the Earth.

Kepler's laws of planetary motion were obtained empirically; they may be stated as follows.

1. The orbit of each planet is an ellipse with the Sun at one focus.
2. The line which joins a planet to the Sun sweeps over equal areas in equal intervals of time.
3. The square of the periods of the planets are proportional to the cubes of their mean distances.

Newton derived these laws from his three laws of motion together with the law of gravitation. They arise from solutions of the *two-body problem* and similar results for spacecraft ensure that such solutions are sufficiently good approximations for many studies of spacecraft navigation (orbit determination and parameter estimation) and guidance (trajectory correction). Any deviations from the basic results of the two-body problem (either because one body is not spherical or because other bodies are present) are called *perturbations*. This chapter is concerned largely with the two-body problem, resulting spacecraft orbits and associated manoeuvres.

1.2 The Two-Body Problem

Consider two masses m_1 and m_2 at positions \mathbf{r}_1 and \mathbf{r}_2 . They are specified to be a large distance apart compared to their size or have spherically symmetrical mass distributions and never touch each other. Therefore they can be treated as particles with the following equations of motion, G being the constant of gravitation.

$$m_1 \ddot{\mathbf{r}}_1 = G m_1 m_2 (\mathbf{r}_2 - \mathbf{r}_1) / r^3 \quad (1.1)$$

$$m_2 \ddot{\mathbf{r}}_2 = G m_1 m_2 (\mathbf{r}_1 - \mathbf{r}_2) / r^3 \quad (1.2)$$

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \quad (1.3)$$

Combining the above equations yields

$$\ddot{\mathbf{r}} + \mu \mathbf{r} / r^3 = 0 \quad (1.4)$$

where

$$\mu = G(m_1 + m_2) \quad (1.5)$$

The centre of mass is at

$$(m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2) / (m_1 + m_2) \quad (1.6)$$

and it is readily confirmed that this position vector is invariant. Relative motion is described by the vector \mathbf{r} but in the common case when one mass is much greater than the other, this becomes virtually the motion of the smaller mass relative to the larger body.

Take the vector product of eqn 1.4 with \mathbf{r} to obtain

$$\frac{d}{dt} (\mathbf{r} \times \frac{d\mathbf{r}}{dt}) = 0 \quad (1.7)$$

The angular momentum per unit mass is

$$\mathbf{h} = \mathbf{r} \times \frac{d\mathbf{r}}{dt} \quad (1.8)$$

and it is constant. Notice that \mathbf{h} is normal to both \mathbf{r} and $\dot{\mathbf{r}}$ and is therefore normal to the plane of motion which must be fixed in space. Now take the vector product of eqn 1.4 with \mathbf{h}

$$\ddot{\mathbf{r}} \times \mathbf{h} = -(\mu/r^3) \mathbf{r} \times (\mathbf{r} \times \dot{\mathbf{r}}). \quad (1.9)$$

Since for any triple vector product

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}) \quad (1.10)$$

it follows that

$$\ddot{\mathbf{r}} \times \mathbf{h} = \mu \frac{d}{dt} (\mathbf{r}/r) \quad (1.11)$$

Since \mathbf{h} is a constant this equation can be integrated directly to yield

$$\dot{\mathbf{r}} \times \mathbf{h} = \mu(\mathbf{r} + r\mathbf{e})/r \quad (1.12)$$

where \mathbf{e} is a constant of integration and is called the eccentricity vector. Notice that when (at closest approach) $\dot{\mathbf{r}}$ is zero, the vector $\dot{\mathbf{r}}$ is perpendicular to \mathbf{r} and $\dot{\mathbf{r}} \times \mathbf{h}$ is along \mathbf{r} . It follows then from eqn 1.12 that \mathbf{e} is along the vector to the point of closest approach. To continue with eqn 1.12 using 1.8

$$\mathbf{r} \cdot \dot{\mathbf{r}} \times \mathbf{h} = \mathbf{r} \times \dot{\mathbf{r}} \cdot \mathbf{h} = h^2 = \mu(r + r e \cos \theta) \quad (1.13)$$

where θ is the angle between the vectors \mathbf{r} and \mathbf{e} , ie. the point of closest approach. Solving for r gives

$$r = \frac{h^2/\mu}{1 + e \cos \theta} \quad (1.14)$$

as the equation of the orbit. Equation 1.14 is the standard equation of a conic in polar form with the origin of coordinates at one focus. The *semilatus rectum* is

$$p = h^2/\mu \quad (1.15)$$

and e is the *eccentricity*. The angle θ is the angle from pericentre (periapsis or for the Earth perigee) or conventionally the *true anomaly*. Referring to eqn 1.8 and Figure 1.1, use the following unit vectors:

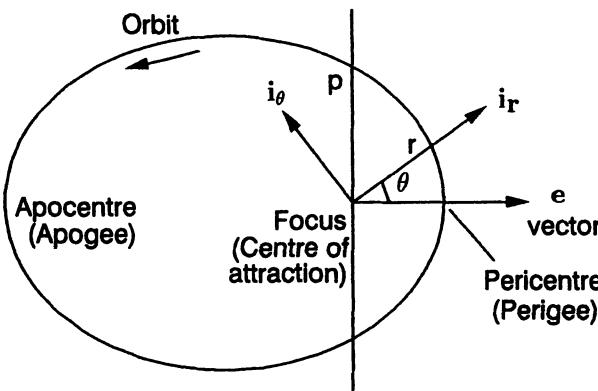


Fig. 1.1. Conic Orbit

1. i_r is along the radius vector \mathbf{r} , ie. away from the centre of attraction.
2. i_θ is perpendicular to \mathbf{r} in the plane of motion in the direction of increasing θ .
3. i_z is the normal to the plane of motion.

Thus eqn 1.8 can be written

$$\mathbf{h} = \mathbf{r} \times (\mathbf{i}_r \dot{r} + \mathbf{i}_\theta r \dot{\theta}) \quad (1.16)$$

or

$$\mathbf{i}_z h = \mathbf{i}_z r^2 \dot{\theta} \quad (1.17)$$

Hence

$$h = r^2 \dot{\theta} \quad (1.18)$$

Eliminate r between the above, eqns 1.14 and 1.15

$$\sqrt{\frac{\mu}{p^3}} dt = \frac{d\theta}{(1 + e \cos \theta)^2} \quad (1.19)$$

Integration of this equation gives θ as a function of time in orbit.

1.3 Orbital Elements and Reference Axes

The parameters p , e and θ represents a choice of the three *orbital elements* necessary to define motion in the plane of the orbit. Before proceeding further let us calculate the energy E per unit mass. By differentiating eqn 1.14 and eliminating $\dot{\theta}$ using eqn 1.18

$$\dot{r} = \mu e \sin \theta / h \quad (1.20)$$

The square of the velocity is given by

$$\begin{aligned} v^2 &= (\dot{r})^2 + (r \dot{\theta})^2 \\ &= h^2/r^2 + (\mu e \sin \theta / h)^2 \\ &= (\mu/h)^2 [1 + e \cos \theta + e^2] \end{aligned} \quad (1.21)$$

Define

$$a = p/(1 - e^2) \quad (1.22)$$

then by using eqns 1.14 and 1.15 the energy E is given by

$$E = v^2/2 - \mu/r = -\mu/(2a) \quad (1.23)$$

In the case of an ellipse, we can readily confirm by substituting θ equal to 0 and π radians in eqn 1.14 with 1.15 that a is the semi-major axis of the ellipse. It is common to employ a , e and θ as the in-plane orbital elements for the hyperbola ($e > 0$) and ellipse ($e < 0$) but a must be regarded as negative for the hyperbola.

In order to specify the orientation of the orbital plane in space three other orbital elements are required. We use the three Euler angles defined in Figure 1.2 as follows:

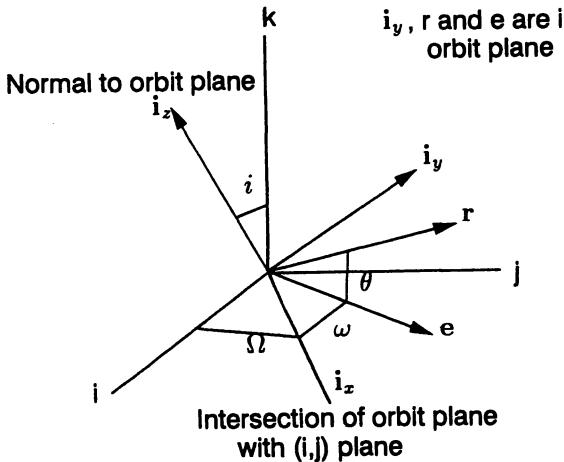


Fig. 1.2. The out-of-plane orbital elements

1. The *longitude of the ascending node* Ω is measured from the x or i axis to the intersection of the orbit plane when the orbit direction is upwards.
2. The *angle of inclination* i is that between the orbit plane and the xy or ij plane.
3. The angle ω is called the *argument of pericentre or periapsis* (perigee or perihelion in the case of the Earth or Sun). It is the angle (measured counter clockwise) from the reference direction i_x of Figure 1.2 to the pericentre through which passes the vector e .

The axes (i, j, k) of Figure 1.2 are intended to be fixed in space, ie. they are inertial axes. They are usually one of two sets as follows:

Celestial or Equatorial Axes. : k is the rotational axis of the Earth (pointing approximately at the Pole star). The axis i is through the point on the celestial equator where the ecliptic (plane of the Earth's orbit about the Sun) crosses the equator going from south to north, ie. the vernal equinox. This was originally pointing at a star the *First Point of Aries* and, although this is sometimes used, the direction of the vernal equinox has drifted slowly away from this constellation. The angular measurements of azimuth and elevation are called *Right Ascension and Declination* respectively.

Ecliptic Axes. : k is normal to the ecliptic plane, ie. the Earth's orbit about the Sun and i again points to the vernal equinox. The angular measurements of azimuth and elevation are called the celestial longitude and latitude respectively.

The calculation of the 6 orbital elements from position and velocity vectors (and vice versa) is now given.

1.3.1 Position and Velocity from the Orbital Elements

Given the 6 orbital elements ($a, e, \theta, i, \Omega, \omega$), by eqns 1.14, 1.15 and 1.22

$$\mathbf{r} = p/(1 + e \cos \theta) \quad (1.24)$$

and \mathbf{v} follows from eqn 1.23. With respect to axes $(\mathbf{i}_r, \mathbf{i}_\theta, \mathbf{i}_z)$ of Figure 1.1, the velocity vector is

$$\begin{aligned} \mathbf{v}_2 &= \dot{r}\mathbf{i}_r + r\dot{\theta}\mathbf{i}_\theta \\ &= \sqrt{\mu/p}[e \sin \theta \mathbf{i}_r + (1 + e \cos \theta)\mathbf{i}_\theta] \end{aligned} \quad (1.25)$$

having used eqns 1.24, 1.20 and 1.18. In these axes

$$\mathbf{r}_2 = r\mathbf{i}_r \quad (1.26)$$

The transformation of the above \mathbf{r}_2 and \mathbf{v}_2 to inertial axes is obtained by two rotational transformations, ie. standard orthogonal transformations. The transformation from axes $(\mathbf{i}_r, \mathbf{i}_\theta, \mathbf{i}_z)$ to axes $(\mathbf{i}_x, \mathbf{i}_y, \mathbf{i}_z)$ is a simple rotation $(\theta + \omega)$, thus

$$\Gamma_1 = \begin{bmatrix} \cos(\theta + \omega) & -\sin(\theta + \omega) & 0 \\ \sin(\theta + \omega) & \cos(\theta + \omega) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1.27)$$

The transformation from axes $(\mathbf{i}_x, \mathbf{i}_y, \mathbf{i}_z)$ to axes $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ is determined by calculating the former three unit vectors with respect to the latter axes. Referring to Figure 1.2

$$\mathbf{i}_x = (\cos \Omega, \sin \Omega, 0) \quad (1.28)$$

$$\mathbf{i}_z = (\sin i \sin \Omega, -\sin i \cos \Omega, \cos i) \quad (1.29)$$

Since $\mathbf{i}_y = \mathbf{i}_z \times \mathbf{i}_x$ it follows that

$$\Gamma_2 = \begin{bmatrix} \cos \Omega & -\sin \Omega \cos i & \sin i \sin \Omega \\ \sin \Omega & \cos \Omega \cos i & -\sin i \cos \Omega \\ 0 & \sin i & \cos i \end{bmatrix} \quad (1.30)$$

The position and velocity vectors can therefore be calculated in inertial axes from

$$\mathbf{r} = \Gamma_2 \Gamma_1 \mathbf{r}_2 \quad (1.31)$$

$$\mathbf{v} = \Gamma_2 \Gamma_1 \mathbf{v}_2 \quad (1.32)$$

1.3.2 Orbital Elements from the Position and Velocity

The position and velocity vectors \mathbf{r} and \mathbf{v} are given with respect to the inertial axes $(\mathbf{i}, \mathbf{j}, \mathbf{k})$. First calculate the in-plane orbital elements. Corresponding to eqn 1.8

$$\mathbf{h} = \mathbf{r} \times \mathbf{v} \quad (1.33)$$

so p is obtained by using the magnitude of \mathbf{h} in eqn 1.15. From the magnitudes of \mathbf{r} and \mathbf{v} we calculate a in eqn 1.23. The eccentricity then follows from eqn 1.22 as

$$e = \sqrt{1 - p/a} \quad (1.34)$$

and from eqn 1.14

$$e \cos \theta = p/r - 1 \quad (1.35)$$

The above does not however permit a unique determination of θ . Note from eqns 1.26 and 1.25 that the scalar product of \mathbf{r} and \mathbf{v} which is independent of axes, is

$$s = \mathbf{r} \cdot \mathbf{v} = r e \sin \theta \sqrt{\mu/p} \quad (1.36)$$

or

$$e \sin \theta = \sqrt{p/\mu s}/r \quad (1.37)$$

and consequently θ can be uniquely determined by

$$\theta = \arctan[s \sqrt{p/\mu}/(p - r)] \quad (1.38)$$

The vector \mathbf{i}_z is calculated as the unit form of $\mathbf{r} \times \mathbf{v}$ and is given at eqn 1.29 in terms of the elements i and Ω . This leads to

$$\Omega = \arctan[\mathbf{i}_z(1)/(-\mathbf{i}_z(2))] \quad (1.39)$$

$$i = \arccos[\mathbf{i}_z(3)] \quad (1.40)$$

The latter does fix i uniquely because it is defined only for positive angles in the range $(0, \pi)$.

In order to derive ω note that the above calculated Ω now determines \mathbf{i}_x by eqn 1.28. Therefore, from Figure 1.2

$$r \cos(\theta + \omega) = \mathbf{i}_x \cdot \mathbf{r} \quad (1.41)$$

Since

$$\mathbf{i}_z r \sin(\theta + \omega) = \mathbf{i}_x \times \mathbf{r} \quad (1.42)$$

$$r \sin(\theta + \omega) = \mathbf{i}_z \cdot (\mathbf{i}_x \times \mathbf{r}) \quad (1.43)$$

Equations 1.41 and 1.43 now yield a unique solution for $(\theta + \omega)$ and hence ω .

At this point it is important to list the problems that can arise in calculating orbital elements due to singularities. They are listed below with suggested remedies.

1. The parabola with $e = 1$. This is unlikely to be a problem in practice, especially if double precision arithmetic is in use. If a parabolic orbit must be handled then the special formulae for this case will have to be introduced; see Battin (Ref. [3]) or Kaplan (Ref. [29]).
2. The circle with $e = 0$. Since the periapsis is not defined ω can be arbitrarily set to any value such as zero. However, see section 3.2.6 for the treatment of near-circular orbits.
3. The inclination i is zero. In this case the line of apsides (\mathbf{i}_x in Figure 1.2) and consequently Ω are undefined. If the inclination is small it is suggested that the axes are rotated through 90 degrees so that the inclination becomes approximately 90 degrees.

1.3.3 Numerical Examples

Some useful data about the solar system is given below.

1. One Astronomical Unit (AU, the semi-major axis of the Earth's orbit about the Sun)= 1.495979×10^8 km
2. One Sidereal Day (spin period of the Earth relative to the stars)= 86164.1 seconds
3. Solar Day (spin period of the Earth relative to the Sun) = 86400 seconds
4. Gravitational constant μ of the Sun = $1.32712 \times 10^{11} \text{ km}^3/\text{sec}^2$
5. Gravitational constant of the Earth = $3.9860 \times 10^5 \text{ km}^3/\text{s}^2$
6. Equatorial (Polar) radius of the Earth = 6378 (6357) km
7. Angle between equatorial and ecliptic planes = 23.45 deg

Table 1.1. Planetary Data

Planet ⁽¹⁾	$\mu 1000\text{km}^3/\text{s}^2$	$a^{(2)}$	Period days	e
Mercury	22.032	0.38710	87.97	0.2056
Venus	324.86	0.72333	224.7	0.00679
Earth	403.50	1.0000	365.26	0.01672
Mars	42.828	1.5237	686.98	0.09338
Jupiter	126712	5.2028	4332.6	0.04845
Saturn	37934	9.5388	10759	0.05565
Uranus	5803.2	19.182	30685	0.04724
Neptune	6871.3	30.058	60189	0.00858
Pluto	40	39.44	90465	0.250
The Moon	4.9028	384401 ⁽³⁾	27.32	0.0549

¹ Including satellites, eg. the Earth and Moon.

² Astronomical Units except the Moon in km.

³ Mean distance because the Sun perturbs the orbit significantly.

The computation of the position and velocity vectors from the orbital elements and vice versa would normally be a computer program, typically

in FORTRAN or C, or for tutorial purposes even by means of a spreadsheet program such as LOTUS123. Therefore the reader might find the following examples useful for checking out such a program, starting from either the orbital elements or position and velocity. The *units* are km, km/s with angles in degrees.

Example 1.3.1. Elliptical Orbit about the Earth :

$$\begin{array}{lll} a = 2.0 \times 10^4 & e = 0.60 & \theta = 60.0 \\ \omega = 170.0 & \Omega = -80.0 & i = 30.0 \\ \mathbf{r} & -7.53 \times 10^3 & 5.10 \times 10^3 & -3.77 \times 10^3 \\ \mathbf{v} & -5.23 & -4.67 & -3.44 \end{array}$$

Example 1.3.2. Hyperbolic Orbit about Jupiter :

$$\begin{array}{lll} a = -9.0 \times 10^4 & e = 2.0 & \theta = 60.0 \\ \omega = 45.0 & \Omega = -80.0 & i = 30.0 \\ M = 13.9 & t_p = 0.302 \text{ hours} & \\ \mathbf{r} & 1.05 \times 10^5 & 5.40 \times 10^4 & 6.52 \times 10^4 \\ \mathbf{v} & 12.39 & 54.54 & 12.51 \\ M = 46.2 & t_p = 0.538 \text{ hours} & \end{array}$$

For an explanation of M and t_p see section 1.4.3

1.4 Time in Orbit

An integral for the time in orbit was given at eqn 1.19 but the solution differs depending on whether the orbit is an ellipse ($e < 1$) or a hyperbola ($e > 1$). The singular case of a parabola ($e = 1$) is ignored here because, although it exists mathematically, an orbit can usually be treated as either an ellipse or a hyperbola.

1.4.1 Solution for Ellipse

In order to simplify the integration the *eccentric anomaly* E is introduced and is defined geometrically in Figure 1.3.

It follows from the geometry of that figure that

$$\cos E = (ae + r \cos \theta)/a \quad (1.44)$$

or

$$r = a(1 - e \cos E) \quad (1.45)$$

This can be equated to the conic equation (from eqns 1.22 and 1.24)

$$r = a(1 - e^2)/(1 + e \cos \theta) \quad (1.46)$$

to yield the following two identities relating the eccentric anomaly to the true anomaly (angle from pericentre or periapsis).

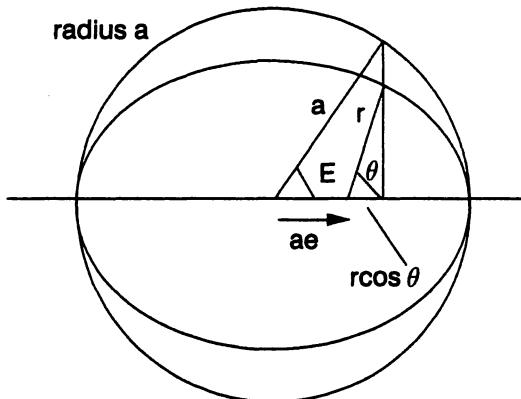


Fig. 1.3. The Eccentric Anomaly

$$\cos \theta = (\cos E - e) / (1 - e \cos E) \quad (1.47)$$

$$\sin \theta = (\sqrt{1 - e^2} \sin E) / (1 - e \cos E) \quad (1.48)$$

Now by differentiation of eqn 1.47

$$\sin \theta d\theta = \frac{(1 - e^2) \sin E dE}{(1 - e \cos E)^2} \quad (1.49)$$

or anticipating the time integral 1.19

$$\frac{(1 - e^2)^{3/2} d\theta}{(1 + e \cos \theta)^2} = (1 - e \cos E) dE \quad (1.50)$$

Integration is now easy and using eqn 1.22 the time from periapsis t_p is related to the eccentric anomaly E by

$$\sqrt{\mu/a^3} t_p = E - e \sin E \quad (1.51)$$

In order to express the angle from periapsis θ in terms of the eccentric anomaly E we use a standard identity for the tangent of a half angle, viz.

$$\tan \frac{\theta}{2} = \frac{1 - \cos \theta}{\sin \theta} \quad (1.52)$$

A little manipulation then permits

$$\tan \frac{\theta}{2} = \sqrt{\frac{1+e}{1-e}} \tan \frac{E}{2} \quad (1.53)$$

Equations 1.51 and 1.53 can be employed to calculate time in orbit given the angle from periapsis, or vice versa. The latter requires however a numerical iterative solution. The angle on the left-hand side of eqn 1.51 is the *mean anomaly* and the period of the ellipse is

$$T = 2\pi \sqrt{a^3/\mu} \quad (1.54)$$

1.4.2 Solution for Hyperbola

The derivation for a hyperbola is similar to that for an ellipse. First introduce a parameter H defined by

$$\cosh H = \frac{a - r}{ae} \quad (1.55)$$

or

$$r = a(1 - e \cosh H) \quad (1.56)$$

Equate this to eqn 1.46 to obtain

$$\cos \theta = \frac{e - \cosh H}{e \cosh H - 1} \quad (1.57)$$

from which follows

$$\sin \theta = \frac{\sqrt{e^2 - 1} \sinh H}{e \cosh H - 1} \quad (1.58)$$

By means of analogous steps to those for the ellipse, the time integral 1.19 is transformed such that integration becomes simple, with the result

$$\sqrt{\frac{\mu}{(-a)^3}} t_p = e \sinh H - H \quad (1.59)$$

The parameter a is of course negative for a hyperbola. Use of the standard identity 1.52 in conjunction with the above expressions for $\sin \theta$ and $\cos \theta$ yields

$$\tan \frac{\theta}{2} = \sqrt{\frac{e + 1}{e - 1}} \tanh \frac{H}{2} \quad (1.60)$$

Equations 1.59 and 1.60 permit the calculation of time in orbit given the angle from periapsis or vice versa, but again a numerical iterative method is necessary for the latter. The left-hand side of eqn 1.59 is again referred to as the *mean anomaly*. Note that since the asymptotes of the hyperbola lie at

$$\theta = \pm \arccos(-1/e) \quad (1.61)$$

there is no solution for θ outside this range.

1.4.3 Computations starting from Time in Orbit

Given the time from periapsis t_p , the calculation of E in eqn 1.51 for an ellipse, or H in eqn 1.21 for a hyperbola cannot be done without resorting to a numerical procedure for solving the equation $f(x) = 0$. Experience shows that it is perfectly satisfactory to employ either the Secant Method

$$x_{i+1} = x_i - \frac{f(x_i)(x_i - x_{i-1})}{f(x_i) - f(x_{i-1})} \quad (1.62)$$

or the Newton-Raphson procedure

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} \quad (1.63)$$

where subscripts denote iteration number.

Given the speed of convergence, it is necessary to initiate the iterations only with values that ensure convergence and the following simple rules will suffice. M is the mean anomaly and it is assumed that the calculations are restricted (without loss of generality) to positive values of M.

- (a) Ellipse $E = \pi$
- (b) Hyperbola $H = 1$ for $M < 2.718e/2$; $H = \log(2M/e)$ for $M \geq 2.718e/2$
However, in calculating θ from E in the case of the ellipse, the following approximation is recommended when E is very close to π

$$\theta = \pi + (E - \pi) \sqrt{\frac{1-e}{1+e}} \quad (1.64)$$

As examples of such calculations the values of M and t_p (corresponding to θ) are listed in the two examples at the end of section 1.3.3.

1.5 Lambert's Time of Flight Theorem

If an orbital transfer occurs from position vector \mathbf{r}_1 to \mathbf{r}_2 in time t then the orbit in a spherical gravitational field is determined. Referring to Figure 1.4, calculate the angle ξ by the cosine formula

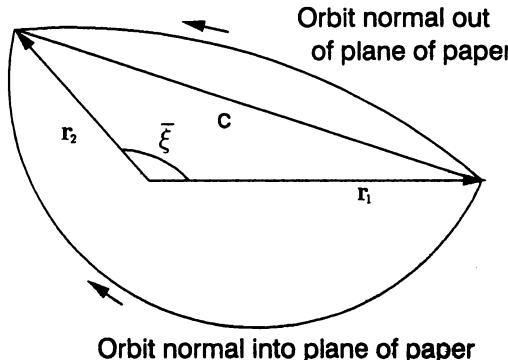


Fig. 1.4. Notation of Orbital Transfers

$$\xi = \arccos \frac{c^2 - r_1^2 - r_2^2}{2r_1r_2} \quad (1.65)$$

and define two cases

1. $\xi = \bar{\xi}$, ie. $\xi \leq \pi$ and the normal to the orbit plane is $\mathbf{r}_1 \times \mathbf{r}_2$ out of the plane of the paper.
2. $\xi = 2\pi - \bar{\xi}$, ie. $\xi > \pi$ and the normal to the orbit plane is $-\mathbf{r}_1 \times \mathbf{r}_2$ into the plane of the paper.

Lambert's Theorem states that the transfer time is a function only of the semi-major axis a , the chord c and $(r_1 + r_2)$, ie.

$$t = F(a, r_1 + r_2, c) \quad (1.66)$$

This theorem complements the orbital algorithms of section 1.3.

If eqn 1.66 is to be used in the forward direction to calculate transfer time then use of the equations below is straightforward. A more common application would however be to calculate the orbit to pass from \mathbf{r}_1 to \mathbf{r}_2 in a given time t . In this case, not only must we resort to numerical iterative computations, but great care to achieve convergence is necessary. Equation 1.66 is nonlinear in a but to apply the Secant or Newton's Method we require a starting approximation in a monotonic region. As it stands although not yet detailed, eqn 1.66 is complicated by (a) an asymptote as a tends to infinity from either positive or negative values, and (b) a cusp associated with two values of transfer time for a given elliptical semi-major axis.

In order to achieve a monotonic function, a transformation from an auxiliary variable u to a is introduced as follows.

$$a = 3(r_1 + r_2 + c)u/8 ; u > 1 \quad (1.67)$$

$$a = (1/4)(r_1 + r_2 + c)/(1 - u^2/3) ; u \leq 1 \quad (1.68)$$

The value of u is then sufficient to define 6 orbital cases (in the notation of Ref. [29], section 7.4) as follows.

u less than $-\sqrt{3}$

1. Hyperbola H1 with $\xi < \pi$
2. Hyperbola H2 with $\xi \geq \pi$

u greater than $-\sqrt{3}$

1. Ellipse E1 excluding apoapsis ($u < 0$) with $\xi < \pi$
2. Ellipse E2 including apoapsis ($u \geq 0$) with $\xi < \pi$
3. Ellipse E3 including apoapsis ($u \geq 0$) with $\xi \geq \pi$
4. Ellipse E4 excluding apoapsis ($u < 0$) with $\xi \geq \pi$

The remaining details of the calculation are now given below without proof (Ref. [29], section 7.4).

$$c = \sqrt{r_1^2 + r_2^2 - 2r_1r_2 \cos \xi} \quad (1.69)$$

$$s = (r_1 + r_2 + c)/2 \quad (1.70)$$

$$\tau_2 = \sqrt{|a|^3/\mu} ; \tau = 2\pi\tau_2 \quad (1.71)$$

$$w_1 = \sqrt{\frac{s}{2|a|}} ; \quad w_2 = \sqrt{\frac{s-c}{2|a|}} \quad (1.72)$$

(a) Hyperbola

$$\gamma = 2 \log \left(w_1 + \sqrt{1 + w_1^2} \right) \quad (1.73)$$

$$\delta = 2 \log \left(w_2 + \sqrt{1 + w_2^2} \right) \quad (1.74)$$

$$v_1 = \sinh \gamma - \gamma \quad (1.75)$$

$$v_2 = \sinh \delta - \delta \quad (1.76)$$

Case H1

$$t = \tau_2(v_1 - v_2) \quad (1.77)$$

$$z = (\gamma + \delta)/2 \quad (1.78)$$

Case H2

$$t = \tau_2(v_1 + v_2) \quad (1.79)$$

$$z = (\gamma - \delta)/2 \quad (1.80)$$

and the semi-latus rectum is given by

$$p = -4a(s - r_1)(s - r_2) \left(\frac{\sinh z}{c} \right)^2 \quad (1.81)$$

(b) Ellipse

Case E1

$$t = \tau_2(u_1 - u_2) \quad (1.82)$$

$$z = (\alpha + \beta)/2 \quad (1.83)$$

Case E2

$$t = \tau - \tau_2(u_1 + u_2) \quad (1.84)$$

$$z = (\alpha - \beta)/2 \quad (1.85)$$

Case E3

$$t = \tau - \tau_2(u_1 - u_2) \quad (1.86)$$

$$z = (\alpha + \beta)/2 \quad (1.87)$$

Case E4

$$t = \tau_2(u_1 + u_2) \quad (1.88)$$

$$z = (\alpha - \beta)/2 \quad (1.89)$$

and the semi-latus rectum is given by

$$p = +4a(s - r_1)(s - r_2) \left(\frac{\sin z}{c} \right)^2 \quad (1.90)$$

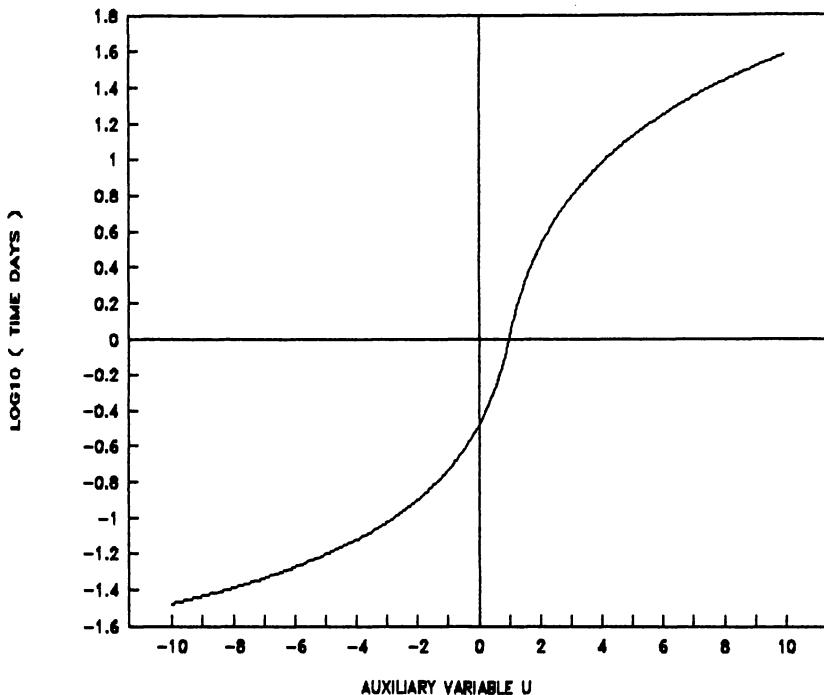


Fig. 1.5. Transfer Time as a Function of u

As an example of the use of this theorem and the value of the above u -to- a transformation, Figure 1.5 shows the transfer time as a function of u for the following numerical data. Given the transfer time for orbits about this comet nucleus, u can be computed efficiently (using the above equations) by means of iterations of the Secant Method (eqn 1.62). In this case, *due to the point of inflexion* evident in Figure 1.5, the Secant Method is much more robust than Newton's Method with respect to starting values.

$$r_1 = 4.91\text{km} \quad r_2 = 7.09\text{km} \quad \xi = 122.5\text{deg} \quad \mu = 1.629 \times 10^4 \text{km}^3/\text{s}^2$$

1.6 Orbital Manoeuvres

Orbital calculations have been so far restricted to the two-body problem in the absence of any propulsive manoeuvres. The transition from one orbit to another by the application of an increment of vector velocity is now analyzed. The results are only approximate with respect to spacecraft trajectories in practice because the propulsive velocity increment is regarded as an impulse, ie. over a period of time short compared to orbital times. The continued restriction to the two-body problem is also a simplification but nonetheless such analysis can be very useful and is widely used in the early stages of mission design.

1.6.1 Coplanar Transfer Manoeuvres

The first example examined is a very common case where propulsive increments are applied to transfer from one circular orbit to another of a different radius but in the same plane. Let the radii of the two circular orbits be r_a and r_p as shown in Figure 1.6. If

$$\sigma = r_a/r_p \quad (1.91)$$

then it follows from eqn 1.24 that

$$e = \frac{\sigma - 1}{\sigma + 1} \quad (1.92)$$

From eqn 1.23 and because $a = (r_a + r_p)/2$

$$r_p v_p^2 / \mu = \frac{2\sigma}{1 + \sigma} \quad (1.93)$$

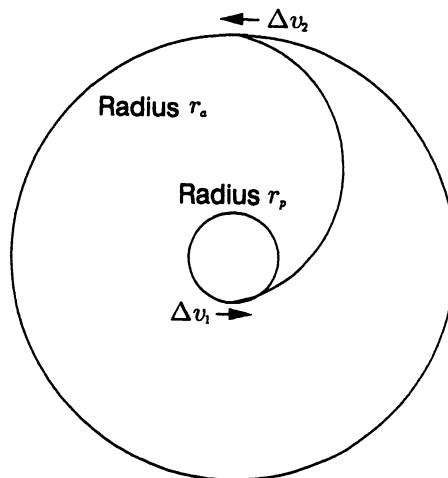


Fig. 1.6. Transfer between Circular Coplanar Orbits

The velocity in the inner circular orbit is $\sqrt{\mu/r_p}$ and therefore the required change of velocity at the periapsis of the transfer ellipse is

$$\Delta v_1 = v_p - \sqrt{\mu/r_p} = \sqrt{\mu/r_p} \left(\sqrt{\frac{2\sigma}{1+\sigma}} - 1 \right) \quad (1.94)$$

To calculate the required velocity increment to enter the outer circular orbit note

$$r_p v_p = r_a v_a \quad (1.95)$$

i.e. by equating the angular momentum h at apoapsis and periapsis. Therefore

$$v_a = (r_p/r_a) \sqrt{\mu/r_p} \sqrt{\frac{2\sigma}{1+\sigma}} \quad (1.96)$$

It follows that the second velocity increment is

$$\Delta v_2 = \sqrt{\mu/r_a} - v_a = \sqrt{\mu/r_a} \left(1 - \sqrt{\frac{2}{1+\sigma}} \right) \quad (1.97)$$

The above analysis can apply to raising or lowering the circular orbit and is known as the *Hohmann Transfer*. It can be shown (Lawden [31]) to be optimum if $\sigma < 11.8$. For higher ratios of circular orbits the use of three impulses is slightly more efficient.

Example 1.6.1. Calculate the velocity increments required to transfer from a low earth orbit (altitude 185 km) to a geostationary circular orbit. The altitude of the latter has a period of one sidereal day (86164.1 seconds from section 1.3.3). Therefore by using the data from the latter section and eqn 1.54 the radius of geostationary orbit is 42164 km. The above formulae then give the first manoeuvre as 2.459 km/s and the second as 1.479 km/s, i.e. a total of 3.938 km/s.

Example 1.6.2. Determine the velocity increment necessary to inject from a circular earth orbit at an altitude of 185 km to an elliptical orbit which just equals the distance of the Moon. The mean distance of the Moon is (section 1.3.3) 384401 km. From the above formula for Δv_1 , the velocity increment is 3.135 km/s, i.e. less than that required for insertion into geostationary orbit.

1.6.2 Injection into an Interplanetary Orbit

Assume that a space vehicle has been placed in a low earth circular orbit as a result of a launch from a site at latitude ϕ with an azimuth angle of α measured clockwise from North. Referring to Figure 1.7 the inclination of the orbit i_0 is most conveniently calculated by spherical trigonometry. Vector algebra (usually preferred by the author) is more ponderous in this case. Note the "right spherical triangle" for which one of the 10 basic identities is immediately what is required, viz.

$$\cos i_0 = \cos \phi \sin \alpha \quad (1.98)$$

Let us use the example of a launch from the NASA site at Cape Kennedy at a latitude of 28.5 degrees. The employed azimuth angles from that site are always in the range 90 to 115 degrees in order to take advantage of the eastward rotation of the Earth and due to restrictions from down-range tracking facilities. Substitution of these numbers into eqn 1.98 shows that the only inclinations possible are in the range 28.5 to 37.2 degrees to the equatorial plane.

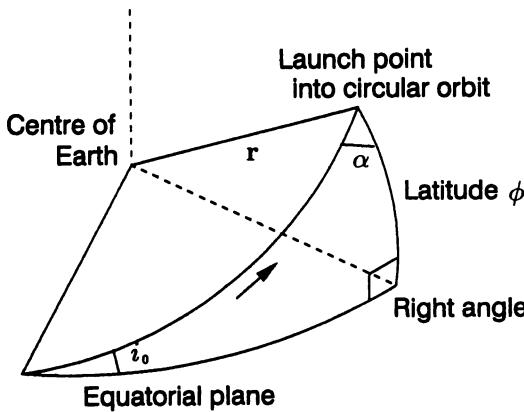


Fig. 1.7. Launch Geometry

Consider now how injection is to be achieved into an interplanetary orbit which is initially a hyperbola relative to the Earth and becomes subsequently an elliptic orbit relative to the Sun, typically in a plane very close to the ecliptic plane. First note as illustrated in Figure 1.8 that, as a result of a short duration boost to velocity v_1 from the circular parking orbit, the velocity vector is turned during escape through the angle ν . From the energy equation 1.23, if v_∞ and r_0 are given

$$v_1 = \sqrt{v_\infty^2 + 2\mu/r_0} \quad (1.99)$$

$$v_\infty^2 = -\mu/a \quad (1.100)$$

$$a = r_0/(1 - e) \quad (1.101)$$

$$e = 1 + r_0 v_\infty^2 / \mu \quad (1.102)$$

$$\theta_\infty = \arccos(-1/e) \quad (1.103)$$

and the angle ν is therefore given by

$$\nu = \arcsin \left(\frac{1}{1 + r_0 v_\infty^2 / \mu} \right) \quad (1.104)$$

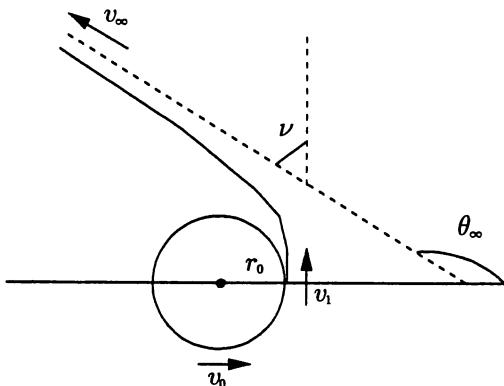


Fig. 1.8. Hyperbolic Escape

Since the velocity in circular orbit is $\sqrt{\mu/r_0}$ the required increment of velocity to escape from circular orbit is

$$\Delta v = \sqrt{v_\infty^2 + 2\mu/r_0} - \sqrt{\mu/r_0} \quad (1.105)$$

e.g. Δv is 4.309 km/s to escape at 5km/s (v_∞) from a parking orbit at an altitude of 185 km. The hyperbolic escape requires no change of orbital plane but, by timing the injection boost from circular orbit and allowing for the turn through angle ν , the hyperbolic escape vector can be pointed with a Declination anywhere between $+i_0$ and $-i_0$. For the example of Cape Kennedy this corresponds to about ± 35 degrees. The ecliptic plane is at 23.45 degrees to the equatorial plane; therefore the *the escape velocity vector can be pointed well away from the ecliptic plane without a plane-change manoeuvre*. However, the inclination of the resulting interplanetary orbit derives from the vector addition of the escape velocity and the velocity of the Earth around the Sun. Nonetheless, missions close to the ecliptic plane can be initiated without serious loss of energy. The Right Ascension of the escape asymptote vector can be adjusted by timing the launch and consequently Ω of the parking orbit.

1.6.3 Plane Change to inject into a Geostationary Orbit

The standard launch site of the European Space Agency (ESA) is at Kourou, a mere 5.5 degrees N of the equator. In this case any parking orbit will be very close to the equatorial plane so that the adjustment to achieve a geostationary orbit at radius 42164 km exactly in the equatorial plane will be very small. On the other hand it has been pointed out that NASA space vehicles launched from Cape Kennedy will have orbital inclinations in the range 28.5 to 37.2 degrees. Unlike hyperbolic escape to planetary missions, achievement of an equatorial geostationary orbit necessitates a plane-change manoeuvre. The

required rotation of the orbital plane is for example 33 degrees. Since the velocity in circular orbit at 185 km is 7.793 km/s a simple procedure would be first to rotate the plane 33 degrees and then carry out the transfer already calculated in Example 1.6.1. This would require

$$7.793 \tan 33 \text{ deg} = 5.061 \text{ km/s} \quad (1.106)$$

A much more efficient procedure is first to inject into the elliptical transfer orbit (Example 1.6.1) by applying the 2.459 km/s *as the circular orbit crosses the equatorial plane*. The perigee and apogee of the elliptical orbit will then both be in the equatorial plane although the inclination of the orbit is still 33 degrees. Since the velocity at apogee is only 1.596 km/s a simple plane-change at apogee would now require only $1.596 \tan 33 \text{ deg} = 1.036 \text{ km/s}$. However we can improve on this by combining the plane change and circularization manoeuvres. Given that a final velocity of 3.075 km/s is required for the circular orbit, the single manoeuvre is obtained by the cosine formula

$$\begin{aligned} \Delta v_2 &= \sqrt{3.075^2 + 1.596^2 - 2 \times 3.075 \times 1.596 \cos 33 \text{ deg}} \\ &= 1.942 \text{ km/s} \end{aligned} \quad (1.107)$$

This is only 0.463 km/s more than the coplanar manoeuvre of 1.479 km/s of Example 1.6.1.

1.7 Patched Conics

The example of section 1.6.2 illustrated how the hyperbolic asymptote of an earth orbit became the start of an elliptical orbit with respect to the Sun. The spacecraft escaped the gravitational influence of the Earth after which the gravitational field of the Sun would become dominant, resulting in an elliptical orbit with respect to the Sun. Trajectories can be approximated in this way, by treating motion along segments as the two-body problem when the gravitational field is due to only one body (along a given segment) such as the Earth, Sun or a planet. The segments are limited by *the spheres of influence* of the major bodies and the approximation is known as the method of patched conics.

1.7.1 Spheres of Influence

Let two major bodies of masses m_1 and m_2 be at a distance R apart. If a spacecraft is distance r_1 from body 1 and distance r_2 from body 2, then the gravitational acceleration of the spacecraft has the same magnitude due to each body when

$$m_1/r_1^2 = m_2/r_2^2 \quad (1.108)$$

If the spacecraft is close to body 1 then $r_2 \approx R$ and therefore the above criterion becomes

$$r_1 = R(m_1/m_2)^{1/2} \quad (1.109)$$

The above equation could be used to give a rough indication of the gravitational sphere of influence of body 1, ie. in terms of the mass ratio of bodies 1 and 2 and their separation distance R.

The concept of the sphere of influence is merely an approximation. Nonetheless a more refined analysis is possible and was first given by Laplace (section 7.1.5, Ref. [29]). Equation 1.109 becomes modified to

$$r_1 = R(m_1/m_2)^{2/5} \quad (1.110)$$

By using the figures in Table 1.1 for mass ratios (from ratios of μ) and distance from the Sun (from a) the following data can be calculated.

Table 1.2. Planetary Spheres of Influence

Planet ⁽¹⁾	Mass Ratio $\times 10^4$ (Planet/Sun)	Radius $\times 10^5$ km of Sphere of Influ.
Mercury	0.00164	1.12
Venus	0.0245	6.16
Earth	0.0304	9.29
Mars	0.00324	5.78
Jupiter	9.55	482
Saturn	2.86	545
Uranus	0.436	519
Neptune	0.518	868
Pluto	0.025	341

1.7.2 Time in Julian Days

Before presenting an example of the use of patched conics it is necessary to introduce a convenient measure of time employed in astronomy and commonly for space missions. The universally adopted system for measuring time in an absolute sense was to define the zero of the time scale as noon January 1st, 4713 BC. The number of days since that starting point is called the *Julian Day*. The history of this curious choice need not concern us; it is sufficient to note for example the Julian Day of a more recent calendar date, eg. noon 4 July 1976 is day number 2442964. We will find it more convenient to use

$$JD1 = JulianDay - 2400000 \quad (1.111)$$

ie. in the above example $JD1 = 42964$.

Sometimes a date might be quoted as say July 4.125 1976 which means 0.125 (3 hours) into 4 July 1976, where the 3 hours refers to a clock at zero longitude, ie. Universal Time (UT) or equivalently Greenwich Mean Time (GMT). Perhaps leading to some confusion is a statement such as "the

Julian Day for Jan 0.0 1995 is $JD1 = 49717.5$ " (Appendix J, Ref. [57]. This is actually referring to a time 24 hours before the time of New Years Eve celebrations and consequently it might be more easily understood if we stated that "noon 1 Jan 1995 is $JD1 = 49719$. A very convenient algorithm has been given (section 1.4, Ref. [57]) for calculating the Julian Day from the date exploiting the integer arithmetic of say FORTRAN or C. Users of the spreadsheet programs LOTUS123 and Microsoft Works should also note that, in terms of the DATE function of both suites

$$JD1 = @DATE(year, month, day) + 15019 \quad (1.112)$$

Corresponding functions permit the inverse process.

1.7.3 A Solar Polar Mission

A mission called Ulysses was designed to have a scientific spacecraft take measurements over the North and South poles of the Sun. Such observations require the spacecraft to be in an orbit almost at 90 degrees to the ecliptic. However, interplanetary orbits are invariably close to the ecliptic plane as a result of the component of the Earth's velocity around the Sun. The result is that a large propulsive correction would be needed to rotate the plane of the orbit through nearly 90 degrees. The same result was achieved in the Ulysses mission by a close approach to Jupiter.

In this example the calculations have been simplified slightly by approximating the motion of the Earth and Jupiter to be in circular orbits in the ecliptic plane. The planetary longitudes at given dates ($\theta + \omega + \Omega$) measured from the vernal equinox are known from tables. For example on 0 Jan 1970 (same as 31 Dec 1969) corresponding to $JD1=40587$, the longitudes of the Earth and Jupiter were 99.74 and 203.42 degrees respectively (Table L-1, Ref. [57]). See Reference [39] for more complete data. Subsequent angles of longitude can therefore be calculated by using the orbital periods quoted in Table 1.1. The resulting data for critical dates of the mission (Ref. [34]) is summarized below. The different phases of the Ulysses mission are illustrated

Table 1.3. Critical Dates of the Ulysses Mission

Event	Date	JD1	Long. Earth	Long. Jupiter
Launch	6 Oct 90	48171	14.6	113.6
Jupiter flyby	8 Feb 92	48661	137.6	154.3
1st polar pass	13 Sept 94	49609	351.9	233.1
Perihelion	12 Mar 95	49789	169.3	248.0
2nd polar pass	31 Jly 95	49930	308.3	259.7

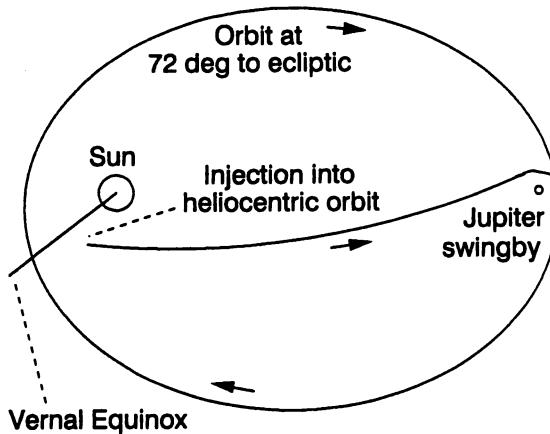


Fig. 1.9. The Solar Polar Mission

in Figure 1.9 and the three distinct orbits are now evaluated.

Transfer from the Earth to Jupiter

Let the planetary longitude of the Earth at injection be α and the hyperbolic escape velocity (relative to the Earth and assumed to be tangential to its orbit) be v_∞ . Then the position and velocity of the spacecraft at injection in ecliptic coordinates are

$$\mathbf{r} = (\cos \alpha, \sin \alpha, 0) 1.0 \text{ A.U.} \quad (1.113)$$

$$\mathbf{v} = (-\sin \alpha, \cos \alpha, 0) (v_0 + v_\infty) \quad (1.114)$$

where v_0 is the orbital velocity of the Earth. When approximated by a circular orbit the latter is 29.785 km/s.

From Table 1.3 the Jupiter flyby should occur 490 days after injection when the longitude of the spacecraft is equal to that of Jupiter, viz. 154.3 degrees. Since injection corresponds to the perihelion of the spacecraft orbit because the velocity is tangential, it follows that for the transfer orbit at flyby $\theta = 154.3 - 14.6$ degrees and r should be equal to a for Jupiter (5.20 A.U.) and t_p equal to 490 days. Thus α (timing of the launch) and v_∞ (from last stage boost) have to be adjusted so that, r and t_p , satisfy these values. Given \mathbf{r} and \mathbf{v} the orbital elements are of course determined by the algorithms of section 1.3.2. Trial and error with computer versions of these algorithms yields $\alpha = 14.8$ degrees and $v_\infty = 10.7$ km/s when $e = 0.848$ and the aphelion of the transfer orbit is at 12.1 A.U.

Jupiter Flyby

Once the spacecraft is within the Jupiter sphere of influence (482×10^5 km from Table 1.2) the orbit can be approximated as a hyperbola with respect to Jupiter. Let \mathbf{v}_1 and \mathbf{v}_2 denote the spacecraft velocity in ecliptic axes before and after the Jupiter flyby. If \mathbf{v}_J is the corresponding velocity of Jupiter, the entry and escape hyperbolic velocities relative to Jupiter are

$$\mathbf{u}_\infty = \mathbf{v}_1 - \mathbf{v}_J \quad (1.115)$$

$$\mathbf{w}_\infty = \mathbf{v}_2 - \mathbf{v}_J \quad (1.116)$$

where

$$\mathbf{v}_J = 13.064(-\sin \alpha_J, \cos \alpha_J, 0) \text{ km/s} \quad (1.117)$$

and α_J is equal to 154.3 degrees from Table 1.3.

Let us examine if \mathbf{v}_2 can have a specified magnitude and perpendicular to the radius vector at Jupiter, viz. \mathbf{r}_J . In this case that point will be the aphelion of the post-Jupiter orbit relative to the Sun and, if \mathbf{n} is the unit vector along \mathbf{v}_J ,

$$\mathbf{v}_2 = v_2(-\sin i \mathbf{k} + \cos i \mathbf{n}) \quad (1.118)$$

\mathbf{k} being the unit normal perpendicular to the ecliptic plane. Thus, if v_2 is given, the inclination i of the third orbit follows by equating u_∞^2 and w_∞^2 from eqns 1.115 and 1.116

$$i = \arccos \left(\frac{v_2^2 + v_J^2 - u_\infty^2}{2v_2v_J} \right) \quad (1.119)$$

Anticipating a value required below, let $v_2 = 8.261 \text{ km/s}$, then since $v_J = 13.064 \text{ km/s}$ we obtain the inclination i equal to 72.3 degrees.

Returning to the hyperbolic orbit around Jupiter, note that the angle between the two equal magnitude vectors \mathbf{u}_∞ and \mathbf{w}_∞ is 2ν ; see Figure 1.8. Therefore

$$\cos(2\nu) = \frac{\mathbf{u}_\infty \cdot \mathbf{w}_\infty}{u_\infty^2} \quad (1.120)$$

and ν can be calculated because both v_1 and v_2 are now known. In fact $2\nu = 71.3$ degrees. Since $\theta_\infty = \nu + \pi/2$ the eccentricity is obtained from

$$e = -1/\cos \theta_\infty = -1/\sin \nu = 1.716 \quad (1.121)$$

Finally the periapsis (closest approach to Jupiter) is calculated from eqn 1.102 as

$$r_0 = \mu(e-1)/u_\infty^2 = 7.337 R_J \quad (1.122)$$

The equatorial radius of Jupiter R_J is 71398 km. *Thus, by means of a relatively close (but safe) approach to Jupiter, the plane of the orbit has been swung through 72.3 degrees without any propulsive expenditure.* Note that the angle turned through relative to ecliptic axes is not 2ν . It is obtained from the scalar product of \mathbf{v}_1 and \mathbf{v}_2 as 80.5 degrees.

High Inclination Solar Orbit

After escaping from the gravitational influence of Jupiter the position vector is still essentially equal to that of Jupiter, viz.

$$\mathbf{r}_J = 5.20(\cos \alpha_J, \sin \alpha_J, 0) \text{ A.U.} \quad (1.123)$$

and the velocity vector is \mathbf{v}_2 which is now known. Therefore a computer program of the formulae of section 1.3.2 can be employed to generate the orbital elements of the third orbit. The results are as follows: $a = 3.25\text{A.U.}$ $e = 0.600$ $\theta = 180\text{deg}$ $\omega = 0.0\text{deg}$ $\Omega = -25.7\text{deg}$ $i = 72.3\text{deg}$

The perihelion is an acceptable 1.30 A.U. and *the inclination of the orbit is sufficiently high to permit observation of the N and S poles of the Sun.*

1.8 Numerical Integration of Orbits

Attention has so far concentrated on the two-body problem and especially when the mass of one body (the spacecraft) is negligible. While this is a powerful tool for approximate mission analysis it is also necessary to be able to compute accurate trajectories when the motion of a spacecraft is subject to the influence of more than one gravitational field or the dominant field is non-spherical. The following summary deliberately ignores classical theory of orbital perturbations, partly because some attention is given to this subject in chapter 3 on earth satellites but also to reflect the availability of computer power undreamed of by mathematicians of the nineteenth century.

Since one gravitational field is usually dominant it is convenient to discuss numerical methods with respect to the following formulation, where \mathbf{r} is the position vector of a spacecraft in inertial axes centred at the dominant body and \mathbf{f} is the sum of all other perturbing forces. The latter are not necessarily small; they may be generated analytically or from computer-stored tables, eg. the ephemerides of the planets.

$$\ddot{\mathbf{r}} + \mu \mathbf{r}/r^3 = \mathbf{f} \quad (1.124)$$

This set of 3 second order differential equations is transformed to 6 first order differential equations by the substitutions

$$\dot{\mathbf{r}} = \mathbf{v} \quad \dot{\mathbf{v}} = \mathbf{f} - \mu \mathbf{r}/r^3 \quad (1.125)$$

1.8.1 Cowell's Method

The numerical integration of eqns 1.125 was first used by Cowell to determine the orbit of a moon of Jupiter and to predict the return of Halley's comet. Even though it is an obvious approach to computer oriented engineers and mathematicians at the end of the twentieth century, it is still often referred to as Cowell's Method.

The coordinates of eqn 1.125 are invariably taken to be rectangular, in equatorial or ecliptic axes and centred at the principal gravitational body. One of several methods of numerical integration can be employed, for example from libraries such as NAG or IMSL. For most purposes when the accuracy demanded is not too high, the fourth order Runge-Kutta process is recommended. If the equation to be integrated is

$$\dot{x} = f(x, t) \quad (1.126)$$

then one step of this Runge-Kutta process is defined by

$$\begin{aligned} k_0 &= hf(x(t_0), t_0) \\ k_1 &= hf(x(t_0) + k_0/2, t_0 + h/2) \\ k_2 &= hf(x(t_0) + k_1/2, t_0 + h/2) \\ k_3 &= hf(x(t_0) + k_2, t_0 + h) \\ x(t_0 + h) &= x(t_0) + (k_0 + 2k_1 + 2k_2 + k_3)/6 \end{aligned} \quad (1.127)$$

from which it is evident that four function evaluations per step are required. *The error per step is the order of h^5* , thus permitting a larger interval of integration (for a given accuracy) than simpler methods. The method is applied to the simultaneous integration of sets of first order differential equations by simply interpreting x, f, k_0, k_1, k_2, k_3 as vectors.

When very high accuracy is required such as before operational missions, then the eighth order Runge-Kutta process is used, again with respect to rectangular coordinates. In this case planetary perturbations are taken into account by means of computer tables of the ephemerides made available by the NASA Jet Propulsion Laboratory (Ref. [50]) and even some relativistic corrections are included (Ref. [36]). In order to enhance confidence in the computed trajectories, it is common practice (for example at the European Space Operations Centre ESOC) to compare the results with those obtained by a second method as described below.

1.8.2 Encke's Method

This method is appropriate to the situation when the motion of a spacecraft is predominantly influenced by one gravitational body, but perturbations arise from other bodies or due to departures from sphericity of the principal field. An approximate *reference or osculating orbit* (from given initial conditions) is needed which would usually be the conic orbit corresponding to the spherical gravitational field of the main gravitational body.

Let the position vector of this reference orbit be $\mathbf{s}(t)$. The true position vector $\mathbf{r}(t)$ is that of the differential equations 1.124 and the difference between the true and reference vector is

$$\delta\mathbf{r} = \mathbf{r} - \mathbf{s} \quad (1.128)$$

Substitute this into eqn 1.124 to obtain

$$\delta\ddot{\mathbf{r}} = \mathbf{f} + [(1 - s^3/r^3)\mathbf{r} - \delta\mathbf{r}]\mu/s^3 \quad (1.129)$$

However the term $(1 - s^3/r^3)$ is the difference of two nearly equal quantities and requires special attention.

Since

$$r^2 = (s_1 + \delta x)^2 + (s_2 + \delta y)^2 + (s_3 + \delta z)^2 \quad (1.130)$$

we can approximate r^2/s^2 by

$$r^2/s^2 = 1 + 2\epsilon + O(\epsilon^2) \quad (1.131)$$

where

$$\epsilon = (s_1\delta x + s_2\delta y + s_3\delta z)/s^2 \quad (1.132)$$

Therefore

$$(1 - s^3/r^3) = 1 - [1 + 2\epsilon + O(\epsilon^2)]^{-3/2} = 3\epsilon + O(\epsilon^2) \quad (1.133)$$

Insert this expression into eqn 1.129 yielding

$$\delta\ddot{\mathbf{r}} = \mathbf{f} + [3\epsilon\mathbf{r} - \delta\dot{\mathbf{r}}]\mu/s^3 \quad (1.134)$$

Assuming the substitution $\delta\mathbf{v} = \delta\dot{\mathbf{r}}$, these become the 6 first order differential equations (in $\delta x, \delta y, \dots, \delta v_z$) that are integrated numerically, eg. by the fourth order Runge-Kutta process. *The accuracy of Encke's Method derives from the technique of resorting to numerical integration only for the small corrective terms to the reference trajectory.*

When this method is employed the whole trajectory is segmented (a) in order that the true orbit differs only slightly from the reference orbit, and (b) if the spacecraft passes from the dominant influence of one body to another. Integration in any segment is initiated with the computed \mathbf{r} and \mathbf{v} at the end of the previous segment.

1.9 Software

Computer programs in source code C++ associated with this chapter have been deposited in the publishers Internet FTP server as mentioned in the preface. They are grouped in the file ORBIT.CPP and comprise the following routines.

Kepler : (section 1.4) given the mean anomaly and eccentricity, it produces the true anomaly.

Orbel : (section 1.3.2) given the position and velocity vectors, it produces the orbital elements $p, e, \omega, i, \Omega, \theta$.

Otime : (section 1.4) given ω_0, e, θ it produces the time from periapsis.

Orbrv : (section 1.3.1) given the orbital elements $p, e, \omega, i, \Omega, \theta$ it produces the position and velocity vectors.

Orbecl : (sections 1.5 and 1.3.2) given the position vectors and times at two points in an orbit, it produces $p, e, \omega, \omega_0, t_0$ and the transformation from orbit to ecliptic axes.

Lambet : (section 1.5) given $r_1, r_2, (t_2 - t_1)$ and the angle between the two radius vectors, it produces p, e, ω .

Lamb : (section 1.5) given r_1, r_2, θ, c, s it produces p and ω .

Note that some of the above routines require also the supplied file MATRIX.CPP of matrix subroutines.

The reader should also be aware of the wealth of up-to-date information and some software made available by NASA organisations and others on Internet web sites, some examples of which are given below.

<http://www.nasa.gov/> : Main NASA site with numerous links to other sites.

<http://www.jpl.nasa.gov/> : Main site of NASA Jet Propulsion Laboratory (JPL).

http://www.nasa.gov/nasa_subjects/nasa_subjectpage.html :
NASA information by subject.

<http://www.jpl.nasa.gov/mip/planet.html> : JPL Planetary missions.

http://www.pst.stsci.edu/spss/release_notes/356/356.html :

Starting point for installing the software to compute the JPL Planetary Ephemeris.

<http://planets.gsfc.nasa.gov/TYPE/planetkey.html> :
Astronomical Almanac available as hard copy or by E-mail.

CHAPTER 2

THE LAUNCH PHASE

2.1 Introduction

The development of rockets as a means of carrying payloads to high altitudes was initiated in the nineteen thirties by von Braun in Germany and Goddard in the USA. The German work culminated in the V2 rocket of World War 2, a single-stage rocket bomb (ballistic missile). Multi-stage rockets are however required to put any significant payload into at least a low-Earth orbit. In the early stages of space programmes (1957 onwards) all launch vehicles were based on the propulsion systems of American and Soviet ballistic missiles, upper stages being added as necessary. Since 1970 launch systems have been developed specifically for space applications, eg. the European Ariane and the American Space Shuttle. The latter is different in two essential respects because it is manned and reusable after an atmospheric reentry and wing-assisted landing. The first stage of a typical system occurs in the atmosphere when an air breathing jet engine is a feasible alternative to a rocket motor which carries the weight of its own oxidiser. The addition of wings to provide aerodynamic lift renders such options potentially very efficient and consequently they are being studied actively for future space transportation.

The V2 rocket had a simple preset guidance system but the insertion of spacecraft into Earth or interplanetary orbits with sufficient accuracy requires compensation for the effects of winds, variations in propulsion performance, and other errors. Such compensation entails *navigation* (determining position, velocity, and orientation of the vehicle) and *guidance* (correcting the trajectory by modifying the direction of thrust). If the trajectory corrections are pre-computed and independent of any in-flight measurements then the guidance is said to be pre-set or "open-loop", which is commonly the case for the first stage of a launch system. Closed-loop guidance is based on the outputs from the navigation system. Early navigation systems for the launch phase were virtually the same as those developed for aircraft or missile applications, namely stabilized platforms carrying accelerometers and gyros, but the trend is towards use of "strapdown" instruments when the heavy platform is eliminated.

This chapter includes the basics of spacecraft propulsion and staging calculations together with some reference material on the better known launch systems employed to date. The equations of motion in the atmosphere are then developed and the approximate optimization of the ascent trajectory is demonstrated. Laws of in-flight guidance are discussed and their mechanization in terms of hardware is outlined.

2.2 Propulsion

2.2.1 Basics

The thrust F of a rocket motor arises from the expulsion of gases at velocity V_e to the atmosphere and is given by (Ref. [11])

$$F = mV_e + (p_e - p_\infty)A_e \quad (2.1)$$

where m is the propellant mass flow rate, A_e is the nozzle exit area, p_e is the gas pressure at the nozzle exit, and p_∞ is the external ambient pressure. By setting the latter to zero it follows that the thrust in vacuum is

$$F_0 = mV_e + p_e A_e \quad (2.2)$$

Therefore

$$F = F_0 - p_\infty \quad (2.3)$$

i.e. the thrust developed against a surrounding atmosphere is always less than that in vacuum. The *specific impulse* is defined as the ratio of the impulse delivered in a short interval of time δt to the weight of propellant consumed in that time. Thus it is

$$I_{sp} = \frac{F\delta t}{mg\delta t} = \frac{F}{mg} \quad (2.4)$$

where $g = 9.807m/s^2$. An alternative expression for the specific impulse could be in terms of the mean thrust \bar{F} , the total burning time T , and the total mass of propellant M , viz.

$$I_{sp} = \frac{\bar{F}T}{Mg} \quad (2.5)$$

Equating thrust to rate of change of momentum yields

$$M \frac{dV}{dt} = F = -gI_{sp} \frac{dM}{dt} \quad (2.6)$$

hence by integration

$$\begin{aligned} \int_0^{\Delta V} dV &= -gI_{sp} \int_{m_0}^{m_f} \frac{dM}{M} \\ \Delta V &= gI_{sp} \log(m_0/m_f) \end{aligned} \quad (2.7)$$

or the *mass ratio*

$$R = m_0/m_f = \exp(\Delta V/gI_{sp}) \quad (2.8)$$

where m_0 and m_f are the initial and final mass respectively. This is the so-called *rocket equation* of Tsiolkovsky first published in 1903.

Solid rocket propellants have specific impulses in the range 280 to 300 seconds. Those for a liquid monopropellant lie in the range 150 to 225 seconds while typical liquid bipropellants develop 300 to 340 seconds. Higher specific impulses such as 450 seconds can be achieved with liquid cryogenic bipropellants but at the expense of complications. In general solid propellants are simpler, more reliable and cheaper than liquid propellants but have lower performance, and they cannot be shut off, eg. when a given increment in velocity has been achieved. The specific impulse of a chemical rocket is proportional to the square root of the chamber temperature divided by the average molecular weight of the exhaust gases. High performance is consequently obtained by means of a high chamber temperature and low average molecular weight of combustion products. Propulsion systems are of course also employed after the launch phase for (a) orbital manoeuvres or small corrections, (b) spin control, (c) attitude control, and (c) momentum management. Because solid rockets cannot be shut off, they are likely to be favoured only for say an orbital insertion when the errors can be corrected subsequently. Orbital manoeuvres would usually be by means of a bipropellant or monopropellant liquid system but other smaller propulsive thrusts might be either with a monopropellant or cold gas system. The latter are very simple but they have a low specific impulse, eg. 50-80 seconds. Electric ion propulsion has conversely a very high effective specific impulse (2000-6000 seconds) but, due to the low thrust level the order of 10 mN, it has been considered to date only for stationkeeping in satellite orbits or continuous thrusting on missions to distances of several astronomical units.

2.2.2 Staging

In practice launch systems consist of more than one stage and, after burn-out of each stage, the empty structure is separated to avoid accelerating unwanted parts of the system. To analyze multi-stage rockets some definitions are needed in addition to the mass ratios in eqn 2.8. For any one stage, write the initial mass as

$$m_0 = m_s + m_p + m_u \quad (2.9)$$

where (m_s, m_p, m_u) are the masses of the structure, propellant, and useful mass respectively. The *structural efficiency* is defined as

$$\epsilon = m_s/(m_s + m_p) \quad (2.10)$$

and the *payload (or useful mass) ratio* is

$$x = m_u/m_0 \quad (2.11)$$

From these definitions and eqn 2.8, since

$$1/R = (m_s + m_u)/m_0 = m_s/m_0 + x \quad (2.12)$$

and

$$\epsilon(1 - x) = m_s/m_0 \quad (2.13)$$

it follows that

$$1/R = x + \epsilon(1 - x) \quad (2.14)$$

A subscript i is now employed to denote the i th stage and note that $m_u(i) = m_0(i + 1)$. The *total payload ratio* can then be calculated as

$$x_{tot} = \frac{m_u(N)}{m_0(1)} = x_1 x_2 \cdots x_N \quad (2.15)$$

From eqn 2.8 the final velocity is

$$V_b(N) = \sum_{i=1}^N c_i \log R_i \quad (2.16)$$

where $c_i = gI_{sp}(i)$. Substitute eqn 2.14 to give the form

$$V_b = - \sum_{i=1}^N c_i \log(x_i + \epsilon_i(1 - x_i)) \quad (2.17)$$

In order to optimize the choice of stages it is proposed to seek the solution of the problem of maximizing the total payload ratio x_{tot} , subject to the constraint of achieving a specified final velocity V_b by choice of the individual payload ratios ($x_1, x_2 \dots x_N$). This is carried out by adjoining the equality constraint 2.17 to the function to be maximized 2.15 by means of a Lagrangian multiplier λ . Thus the augmented function to be maximized is

$$V(x_1 \dots x_n) = \sum_{i=1}^N x_i + \lambda \left[\log V_b - \sum_{i=1}^N c_i \log(x_i + \epsilon_i(1 - x_i)) \right] \quad (2.18)$$

The necessary condition for a maximum is

$$\frac{\partial V}{\partial x_i} = 1/x_i - \lambda \frac{c_i(1 - \epsilon_i)}{x_i + \epsilon_i(1 - x_i)} = 0 \quad (2.19)$$

or

$$x_i = \frac{\epsilon_i}{(c_i \lambda - 1)(1 - \epsilon_i)} \quad (2.20)$$

Substitute this into eqn 2.15 to yield

$$x_{tot} = \frac{\epsilon_1 \epsilon_2 \dots \epsilon_N}{(c_1 \lambda - 1)(1 - \epsilon_1) \dots (c_N \lambda - 1)(1 - \epsilon_N)} \quad (2.21)$$

For the purpose of an illustrative numerical example, simplify the equations by setting all stages to have the same exhaust velocity c and the same structural efficiency ϵ . Thus eqn 2.21 becomes

$$x_{tot} = \left[\frac{\epsilon}{(1 - \epsilon)(c\lambda - 1)} \right]^N \quad (2.22)$$

and

$$x_i = x_{tot}^{1/N} \quad (2.23)$$

Consequently from 2.17

$$V_b = -Nc \log \left[\epsilon + x_{tot}^{1/N} (1 - \epsilon) \right] \quad (2.24)$$

or by rearrangement since V_b is specified

$$x_{tot} = \left[\frac{\exp(-V_b/cN) - \epsilon}{1 - \epsilon} \right]^N \quad (2.25)$$

Table 2.1. Illustrative Multi-stage Rocket

Stage columns. number	Total Payload Ratio ⁽³⁾ % circular orbit ⁽¹⁾	geo. transfer orbit ⁽²⁾
1	0.000	0.000
2	3.525	0.696
3	4.357	1.325
4	4.681	1.567

¹ 300 km altitude above the Earth. $V_b = 7.726$ km/s.

² 185 km altitude of the perigee. $V_b = 10.252$ km/s.

³ Calculated for $\epsilon = 0.1$ and $I_{sp} = 300$ seconds.

Table 2.1 shows the results of eqn 2.25 calculated with typical values for the structural efficiency and specific impulse (bipropellant) of 0.1 and 300 seconds respectively. The two cases are for insertion into a low Earth orbit and a geostationary transfer orbit but remember that these are simple idealised calculations with no allowance for aerodynamic drag, etc. Note that *at least two stages must be employed*, but the extra complications of a fourth stage are not usually worthwhile in this illustration.

2.3 Launch Systems

The injection of a spacecraft into an Earth satellite orbit or a trajectory to outer space requires two, three or four rocket stages which constitute the launch system. The principal launchers that have been used in the past or

are still in service are summarised below. See references [58](ch.18), [16] and [38], or for a more graphic experience, visit the Smithsonian Museum of Air and Space, Washington D.C., USA.

Titan : a family of US vehicles developed from the intercontinental missile. A two-stage liquid-fuelled Titan rocket launched the Gemini manned missions. More powerful Titans, with additional stages and strap-on boosters, were used to launch spacecraft to Mars (Viking and Observer) and the outer planets (Voyager). In conjunction with a Centaur upper stage, Titan IV remains an available US expendable rocket with a lifting capability similar that of the Space Shuttle.

Saturn : a US launcher of the late 1960's and 1970's. A two-stage version (Ib) sufficed for the Earth orbits of the Apollo preparatory flights but the three-stage version (V) was necessary for the Apollo missions to put a man on the Moon. Saturn V was one of the world's most powerful rockets with an initial mass and thrust of 2,700 tonnes and 3,400 tonnes weight respectively. It could place 140 tonnes into a low Earth orbit.

Proton : a Soviet space rocket of up to four stages introduced in 1965, but it was not successful for space missions until 1969. It was used to launch heavy satellites, lunar landers, space probes, and the Salyut and Mir space stations. Proton can place about 20 tonnes into a low Earth orbit or send about 5 tonnes to the inner planets.

Space Shuttle : a reusable US crewed vehicle first launched in 1981; it is launched conventionally (vertically upwards) but lands like an aircraft. Four orbiters (the reusable part) were built with the names Columbia, Challenger, Discovery, and Atlantis but, after an accident in 1986, Challenger was replaced by Endeavor in 1992. The orbiter weighs 68 tonnes and can deploy a payload of 23 tonnes from its low Earth orbit. The propulsion units are roughly equivalent to two stages because the tank for the liquid propellant of the three main engines is separated when empty and the two solid-fuel boosters are jettisoned two minutes after launch. For higher orbits the 23 tonne "payload" consists partly of an upper stage, or even multiple payloads with their own upper stages.

Energia : a powerful two-stage Soviet launcher dating from 1987, a fresh design following mixed experience with Proton. It was employed to launch the Soviet space shuttle Buran and, with strap-on-boosters, can put up to 190 tonnes into low Earth orbit. A third stage (included in Buran) is necessary for satellite injection.

Ariane : a series of vehicles (versions 1-5) of the European Space Agency launched from Kourou in French Guiana near the Equator. Version 1 first flew in 1979 as a three-stage vehicle but the performance has been progressively improved and different options of solid or liquid fuel strap-on-boosters are available. Ariane 5 can place 18 tonnes in a low Earth orbit or 6.8 tonnes in geostationary transfer orbit. This programme has been run commercially since 1984 by Arianespace.

Atlas-Centaur : a combination of Atlas (lower stages) and Centaur (upper stage) to provide an unmanned US alternative to the Shuttle. Atlas was the first US intercontinental ballistic missile. With the Agena upper stage it was used in the 1960's to launch Mariner probes to Mars and Venus, and for the Ranger lunar series. Atlas-Centaur is a new version designed for commercial orders in the 1990's. It has a capability for placing 5.7 tonnes in low Earth orbit or 2.2 tonnes in geostationary transfer orbit.

Long March 3 : a Chinese three stage vehicle dating from 1984 that will place 3.0 tonnes in low Earth orbit. The first commercial launch (of the US made Asiasat1) took place in 1990.

H-2 : a two-stage Japanese launch system capable of putting 9.0 tonnes in low Earth orbit. Earlier rockets in the N series were based on the US Thor-Delta, but the H-2 is entirely of Japanese manufacture.

2.4 Trajectory Dynamics

2.4.1 Equations of Motion

In order to compute and optimize launch trajectories the equations of motion must be set up in a convenient frame of reference. The basic vector equation

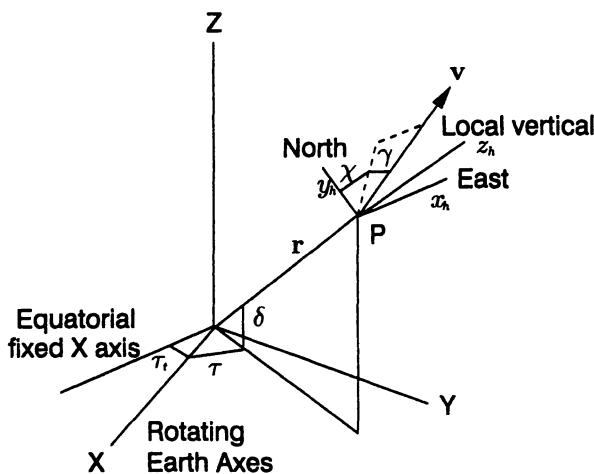


Fig. 2.1. Vehicle-centred horizontal reference axes

of motion is

$$M \frac{d\mathbf{V}}{dt} = \mathbf{F} + \mathbf{W} + \mathbf{A} \quad (2.26)$$

where M and \mathbf{V} are the mass and velocity vector of the rocket at time t . \mathbf{F} is the effective thrust vector of the rocket motor, \mathbf{W} is the gravitational force vector, and \mathbf{A} is the vector of aerodynamic forces. Referring to Figure 2.1, because the XYZ axes are rotating at Earth rate Ω_e

$$\mathbf{V} = \mathbf{v} + \Omega_e \times \mathbf{r} \quad (2.27)$$

where $\mathbf{v} = \dot{\mathbf{r}}$, ie. \mathbf{v} is velocity relative to axes fixed in the rotating Earth. It is the *air velocity* in the absence of winds. Now from eqn 2.27

$$\begin{aligned} \frac{d\mathbf{V}}{dt} &= \frac{d\mathbf{v}}{dt} + \Omega_e \times \mathbf{v} + \Omega_e \times (\mathbf{v} + \Omega_e \times \mathbf{r}) \\ &= \frac{d\mathbf{v}}{dt} + 2\Omega_e \times \mathbf{v} + \Omega_e \times (\Omega_e \times \mathbf{r}) \end{aligned} \quad (2.28)$$

However it is proposed to employ the h axes of Figure 2.1 which are rotating Ω_h relative to the XYZ axes. Therefore

$$\frac{d\mathbf{v}}{dt} = \dot{\mathbf{v}} + \Omega_h \times \mathbf{v} \quad (2.29)$$

giving finally

$$\frac{d\mathbf{V}}{dt} = \dot{\mathbf{v}} + \Omega_h \times \mathbf{v} + 2\Omega_e \times \mathbf{v} + \Omega_e \times (\Omega_e \times \mathbf{r}) \quad (2.30)$$

$$= (\mathbf{F} + \mathbf{W} + \mathbf{A})/M \quad (2.31)$$

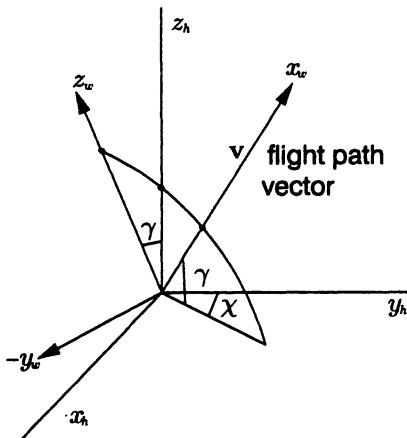


Fig. 2.2. Flight path or wind axes

The components are now evaluated in h axes. First calculate the velocity of point P (Figure 2.2) in h axes.

$$\mathbf{v} = \dot{\mathbf{r}} = \begin{bmatrix} r\dot{\tau} \cos \delta \\ r\dot{\delta} \\ \dot{r} \end{bmatrix} = v \begin{bmatrix} \cos \gamma \sin \chi \\ \cos \gamma \cos \chi \\ \sin \gamma \end{bmatrix} \quad (2.32)$$

It follows that the three equations $\mathbf{v} = \dot{\mathbf{r}}$ become

$$\dot{\tau} = (v/r) \cos \gamma \sin \chi (\cos \delta)^{-1} \quad (2.33)$$

$$\dot{\delta} = (v/r) \cos \gamma \cos \chi \quad (2.34)$$

$$\dot{\gamma} = v \sin \gamma \quad (2.35)$$

In order to evaluate the left hand side of eqn 2.30 note the following in h axes.

$$\mathbf{r} = (0, 0, r) \quad (2.36)$$

$$\boldsymbol{\Omega}_e = \omega_e (0, \cos \delta, \sin \delta) \quad (2.37)$$

$$\begin{aligned} \boldsymbol{\Omega}_h &= (-\dot{\delta}, \dot{\tau} \cos \delta, \dot{\tau} \sin \delta) \\ &= (v/r)(-\cos \gamma \cos \chi, \cos \gamma \sin \chi, \cos \gamma \sin \chi \tan \delta) \end{aligned} \quad (2.38)$$

$$\boldsymbol{\Omega}_e \times \mathbf{v} = \omega_e v \begin{bmatrix} \cos \delta \sin \gamma - \sin \delta \cos \gamma \cos \chi \\ \sin \delta \cos \gamma \sin \chi \\ -\cos \delta \cos \gamma \sin \chi \end{bmatrix} \quad (2.39)$$

$$\boldsymbol{\Omega}_e \times \mathbf{r} = \omega_e r (\cos \delta, 0, 0) \quad (2.40)$$

$$\boldsymbol{\Omega}_e \times (\boldsymbol{\Omega}_e \times \mathbf{r}) = \omega_e^2 r (0, \cos \delta \sin \delta, -\cos^2 \delta) \quad (2.41)$$

By differentiation of eqn 2.32

$$\dot{\mathbf{v}} = \Phi \begin{bmatrix} \dot{v} \\ v \dot{\chi} \cos \gamma \\ v \dot{\gamma} \end{bmatrix} \quad (2.42)$$

where the orthogonal matrix

$$\Phi = \begin{bmatrix} \cos \gamma \sin \chi & \cos \chi & -\sin \gamma \sin \chi \\ \cos \gamma \cos \chi & -\sin \chi & -\sin \gamma \cos \chi \\ \sin \gamma & 0 & \cos \gamma \end{bmatrix} \quad (2.43)$$

From eqns 2.32 and 2.38

$$\boldsymbol{\Omega}_h \times \mathbf{v} = (v^2/r) \cos \gamma \mathbf{s} \quad (2.44)$$

where

$$\mathbf{s} = \begin{bmatrix} \sin \chi (\sin \gamma - \cos \gamma \cos \chi \tan \delta) \\ \sin^2 \chi \cos \gamma \tan \delta + \cos \chi \sin \gamma \\ -\cos \gamma \end{bmatrix} \quad (2.45)$$

The right hand side of eqn 2.30 can now be written down as

$$\Phi \begin{bmatrix} \dot{v} \\ v \cos \gamma \dot{\chi} \\ v \dot{\gamma} \end{bmatrix} + 2\omega_e v \mathbf{q} + \omega_e^2 r \mathbf{p} + (v^2/r) \cos \gamma \mathbf{s} \quad (2.46)$$

where

$$\mathbf{q} = \begin{bmatrix} \cos \delta \sin \gamma - \sin \delta \cos \gamma \cos \chi \\ \sin \delta \cos \gamma \sin \chi \\ -\cos \delta \cos \gamma \sin \chi \end{bmatrix} \quad (2.47)$$

$$\mathbf{p} = (0, \cos \delta \sin \delta, -\cos^2 \delta) \quad (2.48)$$

and \mathbf{s} is given by eqn 2.45.

Now, by evaluating unit vectors in Figure 2.2, it can be confirmed that Φ^T transforms from \mathbf{h} to the flight path or wind axes of that figure, hence introduce the notation

$$\Phi^T(\mathbf{F} + \mathbf{W} + \mathbf{A})/M = (\mathbf{f} + \mathbf{w} + \mathbf{a})/M \quad (2.49)$$

because it is proposed to evaluate the forces in wind axes. Thus operate on eqn 2.46 by Φ^T and use eqn 2.49 to yield

$$\dot{v} = \omega_e^2 r \cos \delta (-\cos \gamma \cos \chi \sin \delta + \sin \gamma \cos \delta) + (f_1 + w_1 + a_1)/M \quad (2.50)$$

$$\begin{aligned} \dot{\chi} &= 2\omega_e (-\tan \gamma \cos \chi \cos \delta + \sin \delta) + \left(\frac{\omega_e^2 r}{v \cos \gamma} \right) \sin \chi \cos \delta \sin \delta \\ &+ (v/r) \cos \gamma \sin \chi \tan \delta + (f_2 + w_2 + a_2)/(Mv \cos \gamma) \end{aligned} \quad (2.51)$$

$$\begin{aligned} \dot{\gamma} &= 2\omega_e \sin \chi \cos \delta + (\omega_e^2 r/v) \cos \delta (\sin \gamma \cos \chi \sin \delta + \cos \gamma \cos \delta) \\ &+ (v/r) \cos \gamma + (f_3 + w_3 + a_3)/(Mv) \end{aligned} \quad (2.52)$$

The components of the forces in wind axes are now calculated.

2.4.2 Gravitational Forces

If $g(r, \delta)$ is the acceleration due to gravity as a function of Earth radius and latitude, the vector of gravitational forces in wind axes (Figure 2.2) is

$$\mathbf{w} = (-g(r, \delta) \sin \gamma, 0, -g(r, \delta) \cos \gamma) M \quad (2.53)$$

The gravitational acceleration in expansion terms up to second order is (section 5.2, Ref. [57])

$$g(r, \delta) = \mu/r^2 [1 + 3J_2(R_0/r)^2(1 - 3\sin^2 \delta)/2] \quad (2.54)$$

where R_0 is the mean equatorial radius. See Table 2.2 below for numerical values. The Earth radius as a function of latitude is given by (section 4.3, Ref. [57])

$$R(\delta) = R_0(1 - f \sin^2 \delta) \quad (2.55)$$

and therefore altitude is

$$h = r - R \quad (2.56)$$

2.4.3 Rocket Thrust

The effective thrust in vacuum of the i'th stage is

$$F_{0i} = I_{sp}(i)b_i \quad (2.57)$$

where $I_{sp}(i)$ is the specific impulse and b_i is the propellant burning rate. It follows that the total mass during the i'th stage is given by

$$M = M_i - b_i(t - t_{0i}) \quad (2.58)$$

where M_i is the mass at the beginning of that stage at time t_{0i} . The thrust in the atmosphere is less due to the effect of the ambient pressure $p(h)$ according to

$$F_i = F_{0i} - A_{nz}p(h) \quad (2.59)$$

where A_{nz} is the nozzle area of that stage. Let the direction of the thrust vector be expressed with respect to the wind axes in terms of the polar angles σ (above the X_wY_w plane) and ϵ (clockwise from the X_w axis). Then the thrust vector during the i'th stage is

$$\mathbf{f} = F_i (\cos \sigma \cos \epsilon, -\cos \sigma \sin \epsilon, \sin \sigma) \quad (2.60)$$

Because the thrust vector must pass through the cg and nominally along the longitudinal X-axis of the rocket, the unit vector along the longitudinal axis is in w axes

$$\tilde{\mathbf{f}} = (\cos \sigma \cos \epsilon, -\cos \sigma \sin \epsilon, \sin \sigma) \quad (2.61)$$

Pointing of the thrust vector is actually mechanized in practice with respect to pitch and yaw angles. Define the longitudinal axis in terms of polar angles in h axes; pitch θ is above the local horizontal X_hY_h plane and yaw ψ is measured clockwise from North, the Y_h axis. Then the longitudinal axis in h axes is $(\cos \theta \sin \psi, \cos \theta \cos \psi, \sin \theta)$. It follows from eqns 2.43 and 2.61 that the angles σ and ϵ can be calculated from the pitch and yaw angles by means of the set

$$\begin{bmatrix} \cos \sigma \cos \epsilon \\ -\cos \sigma \sin \epsilon \\ \sin \sigma \end{bmatrix} = \Phi^T \begin{bmatrix} \cos \theta \sin \psi \\ \cos \theta \cos \psi \\ \sin \theta \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \quad (2.62)$$

Thus

$$\sigma = \arcsin(z_3) \quad \epsilon = \arctan(-z_2/z_1) \quad (2.63)$$

2.4.4 Aerodynamic Forces

The aerodynamic forces are a function of the angle of attack α and the angle of yaw but, assuming symmetry about the longitudinal axis, only the angle of attack is required. If \mathbf{i} denotes the unit vector along the axis X_w (the direction of the air velocity vector), then using eqn 2.61

$$\cos \alpha = \tilde{\mathbf{f}} \cdot \mathbf{i} = \cos \sigma \cos \epsilon \quad (2.64)$$

The normal to the plane containing $\tilde{\mathbf{f}}$ and \mathbf{i} is

$$\mathbf{n} = \tilde{\mathbf{f}} \times \mathbf{i} = (0, \sin \sigma, \cos \sigma \sin \epsilon) \quad (2.65)$$

Therefore the direction of the lift vector (perpendicular to \mathbf{v} or X_w) is obtained from

$$\mathbf{i} \times \mathbf{n} = (0, -\cos \sigma \sin \epsilon, \sin \sigma) / \sin \alpha \quad (2.66)$$

The above unit vector tends to $(0, 0, 1)$ as α tends to zero.

Thus in terms of conventional aerodynamic notation (for example section 14 of Ref. [11]) the lift and drag forces give rise to

$$\mathbf{a} = D(\alpha) \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} + L(\alpha) / \sin \alpha \begin{bmatrix} 0 \\ -\cos \sigma \sin \epsilon \\ \sin \sigma \end{bmatrix} \quad (2.67)$$

where the drag and lift coefficients (for a characteristic area S) are given by

$$D(\alpha) = \rho(h)v^2 C_D(\alpha)S/2 \quad (2.68)$$

$$L(\alpha) = \rho(h)v^2 C_L(\alpha)S/2 \quad (2.69)$$

The drag and lift coefficients (as a function of Mach number) must be determined by elaborate computations and confirmed by wind tunnel measurements. However, if the front of the launch vehicle can be approximated by a cone of semi-angle θ_c , then the following approximate formulae (section 9, Ref. [45]) will usually suffice for preliminary computations. The characteristic area S is the cross sectional area of the base of the cone.

$$\begin{aligned} C_N &= \cos^2 \theta_c \sin 2\alpha \\ C_A &= 2 \sin^2 \theta_c + \sin^2 \alpha (1 - 3 \sin^2 \theta_c) \end{aligned} \quad (2.70)$$

where

$$\begin{aligned} C_D &= +\cos \alpha C_A + \sin \alpha C_N \\ C_L &= -\sin \alpha C_A + \cos \alpha C_N \end{aligned} \quad (2.71)$$

The atmospheric pressure p and density ρ appear in equations above; they must be obtained from models of the atmosphere expressed in tabular form. The 1962 U.S. Standard Atmosphere is for example tabulated by Regan (Table 2.1, Ref. [58]) giving the pressure ratio and temperature as a function of (geometric) altitude. If p_0 is the pressure at sea level put

$$r = \log_e(p/p_0) \quad (2.72)$$

i.e. the log of the pressure ratio which is tabulated at altitudes h_0 to h_{N-1} . Use linear interpolation on tabulated values of r corresponding to exponential interpolation on the pressure ratios. Given the pressure, the density is calculated from

$$\rho = \frac{p}{R_{gas}Temp} \quad (2.73)$$

2.4.5 Final Orbital Elements

It can be assumed that burn-out of the last stage will occur when the vehicle is essentially outside the atmosphere. In that case the elements of the resulting orbit can be calculated from the position and velocity by the formulae specified in section 1.3.2. A computer subroutine can be assumed to exist but the input must be the position and velocity vectors with respect to inertial (non-rotating) axes.

The total velocity \mathbf{V} (including a component from the Earth's rotation) is given at eqn 2.27 and its components in h axes are obtained from eqns 2.32 and 2.40, viz.

$$\mathbf{V} = \begin{bmatrix} v \cos \gamma \sin \chi + \omega_e r \cos \delta \\ v \cos \gamma \cos \chi \\ v \sin \gamma \end{bmatrix} \quad (2.74)$$

The position vector \mathbf{r} in h axes is noted at eqn 2.36. This pair (\mathbf{r}, \mathbf{V}) must be transformed to the non-rotating axes $X_0Y_0Z_0$ of figure 2.1. In that figure let $(\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z)$ be the unit vectors along the h axes (x_h, y_h, z_h) respectively. Note that with respect to the axes $X_0Y_0Z_0$

$$\begin{aligned} \mathbf{e}_x &= (-\sin \tau', \cos \tau', 0) \\ \mathbf{e}_y &= (-\sin \delta \cos \tau', -\sin \delta \sin \tau', \cos \delta) \\ \mathbf{e}_z &= (\cos \delta \cos \tau', \cos \delta \sin \tau', \sin \delta) \end{aligned} \quad (2.75)$$

where

$$\tau' = \tau_t + \tau = \tau_0 + \omega_e t + \tau \quad (2.76)$$

τ_0 being the right ascension at zero time. It follows that the transformation from h to $X_0Y_0Z_0$ axes is

$$\Phi_0 = [\mathbf{e}_x \ \mathbf{e}_y \ \mathbf{e}_z] = \begin{bmatrix} -\sin \tau' & -\sin \delta \cos \tau' & \cos \delta \cos \tau' \\ \cos \tau' & -\sin \delta \sin \tau' & \cos \delta \sin \tau' \\ 0 & \cos \delta & \sin \delta \end{bmatrix} \quad (2.77)$$

This transformation applied to \mathbf{r} and \mathbf{V} in h axes yields those vectors in inertial axes as required to compute the orbital elements.

2.5 Computation of the Nominal Ascent Trajectory

Equations 2.33 to 2.35 and 2.50 to 2.52 constitute the state equations in the dependent variables $\tau, \delta, r, v, \chi, \gamma$. It is required to integrate numerically these equations from given initial conditions while adjusting the thrust vector such that certain final conditions are satisfied. That adjustment of the thrust profile should also maximize the final velocity (or mass) subject to possible side constraints such as the following.

1. lift acceleration which is equal to $\frac{1}{2}\rho v^2 C_L S/M$.
2. instantaneous heat flux which is proportional to $\frac{1}{2}\rho v^3$.
3. total heat which is the integral of the heat flux.
4. range safety limits on latitude (δ) and longitude (τ).

After burn-out of the last stage the vehicle should be in an orbit to satisfy the criteria of the particular mission, eg. a low-Earth orbit possibly to be subsequently modified, a geostationary transfer orbit (GTO), or a parking orbit prior to injection into a hyperbolic escape orbit to outer space. In all such cases the spacecraft will carry a separate propulsion system for the required manoeuvres and need not be considered in these calculations. It is necessary here only to specify the required orbit at termination of the last stage. For example, for GTO it would be required that

1. Apogee $r_{max} = 42164$ km. radius
2. Perigee $r_{min} \approx 300$ km. altitude (to permit one or two revolutions in GTO) without re-entering the atmosphere.
3. Latitude $\delta=0$
4. Path angle from North $\chi = 90$ degrees.

The longitude in the final geostationary orbit would be obtained approximately by timing of the launch and subsequent refinement by means of small corrective manoeuvres in the final operational orbit.

The computation of the best ascent trajectory is theoretically a nonlinear optimization subject to terminal and side constraints, ie. to maximize final mass into orbit subject to the above side and terminal constraints. As such it can be tackled by means of the classical calculus of variations, modern equivalent techniques from optimal control theory (Ref. [8]), or other iterative parametric optimizations (Refs. [7], [56]). In fact *a close approximation to the optimum ascent trajectory can be computed using simple considerations*, provided unusual manoeuvres are not necessary due to range safety limitations arising from an undesirable location of a launch site (Ref. cite36).

The following are those assumptions which enable the computation of a nominal ascent trajectory.

1. A near vertical initial ascent with almost zero angle of attack is necessary to escape the atmosphere as quickly as possible. Loss of energy due to drag is minimized while lateral acceleration due to lift is kept below a tolerable value.
2. Even outside the atmosphere, the thrust vector (along the longitudinal axis) should be close to the velocity vector in order to maximize the rate of increase of energy.
3. Since the final velocity is invariably to be close to the local horizontal, there must be a transition from the near vertical initial ascent to such a final trajectory.

The latter is accomplished by the so-called *gravity turn* and some elementary analysis is instructive in understanding this feature of an ascent.

In order to achieve an approximate analytical solution, apply the following simplifications to equations 2.50 and 2.52:

1. thrusting along the velocity vector so that σ and ϵ are zero,
2. a flat non-rotating Earth so that g is constant and terms involving ω_e are ignored.
3. no atmosphere and hence no aerodynamic terms.

By setting $F/M = g\beta_0$ with $\beta_0 > 1$ equations 2.50 and 2.52 become

$$\dot{v} = g(\beta_0 - \sin \gamma) \quad (2.78)$$

$$\dot{\gamma} = -(g/v) \cos \gamma \quad (2.79)$$

Eliminate time t by dividing the two equations and integrate to yield

$$\log_e v = -\log_e \cos \gamma - \beta_0 \log \left(\frac{1 + \tan \gamma/2}{1 - \tan \gamma/2} \right) + \text{constant} \quad (2.80)$$

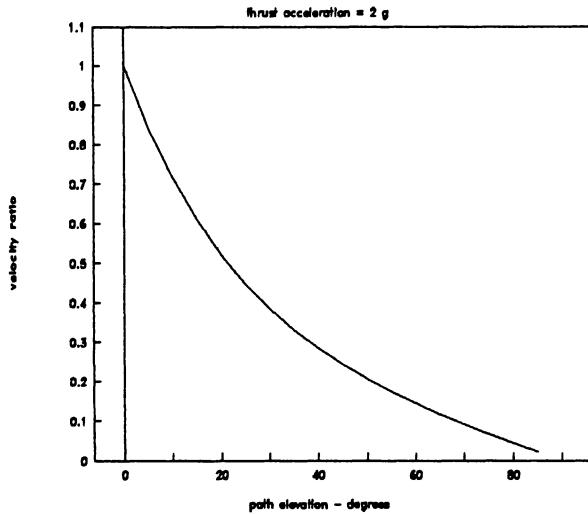


Fig. 2.3. The Gravity Turn

Since

$$\left(\frac{1 + \sin \gamma}{1 - \sin \gamma} \right) = \left(\frac{1 + \tan \gamma/2}{1 - \tan \gamma/2} \right)^2 \quad (2.81)$$

the solution can be written as

$$v \cos \gamma \left(\frac{1 + \sin \gamma}{1 - \sin \gamma} \right)^{\beta_0/2} = \text{constant} \quad (2.82)$$

This function is plotted in Figure 2.3 with the constant set equal to one, ie. corresponding merely to velocity in horizontal flight scaled to unity. Thus

it is seen that, with a very small initial perturbation away from a vertical ascent, the path angle bends over naturally under the influence of gravity as the rocket is accelerated. In practice, after a short interval of vertical ascent, a small tip-over manoeuvre is executed after which the gravity turn develops naturally.

To summarize, near optimum ascent trajectories can be generated by simulating a small tip-over shortly after lift-off, followed by thrusting with all stages (a) at zero angle of attack in the atmosphere, and (b) along the velocity vector outside the atmosphere. In fact the total velocity vector \mathbf{V} is not coincident with the *air velocity vector* \mathbf{v} but the angle is usually small enough to be of little concern in the above statement, eg. a worst case of about 9 degrees just after leaving the atmosphere in a N or S heading ascent.

2.6 A Computed Example: Ariane 4

First the values of a number of constants introduced above are listed in Table 2.2.

Table 2.2. Numerical Data for Ascent Trajectories

Constant	Symbol	Numerical value
air density at sea level	ρ_0	$1.2250 \times 10^9 \text{ kg/km}^3$
gas constant	R	287 joules/kg-deg
grav. const. of Earth	μ	$3.9860 \times 10^5 \text{ km}^3/\text{s}^2$
equatorial radius	R_0	6378.14
polar radius	R_p	6356.75
grav. oblateness coef.	J_2	0.0010826
Earth rotation rate	ω_e	$7.29211 \times 10^{-5} \text{ rad/s}$
pressure at sea level	p_0	$1.013 \times 10^5 \text{ N/m}^2$

Ariane is the launch system for most missions of the European Space Agency (ESA). This computed example is for a spacecraft to be placed in geostationary orbit (GTO) and consequently the required final conditions are as listed above. Since the launch site is so close to the equator (at Kourou), the need for a very small change of the orbital plane is ignored. Although the numerical data has been taken from an ESA publication (Ref. [37]), the simulation can be regarded only as approximate, either due to simplifications or estimates of the author necessitated by the absence of complete data. The case of a single-payload launch of the Ariane 4 (version 40 without boosters) has been chosen and the relevant data is summarised in Table 2.3.

The rocket nozzle areas for the three stages have been estimated as 2.96, 0.78, and 0.06 sq.m respectively. The vehicle cross-sectional area was taken

Table 2.3. Representative Stage Data for Ariane 40

Stage	Structure tonnes	Propellant tonnes	Vacuum Thrust kN	Burn time ⁽²⁾ seconds
1	17.5	157.0 ⁽³⁾	4x748	138
2	4.325 ⁽¹⁾	34.0	760 ⁽⁴⁾	130
3	1.200	10.7	62	735

¹ This includes the 725 kg fairing, equivalent to dropping it at the end (instead of the middle) of the second stage.

² Few seconds in between stages has been omitted.

³ More propellant is employed in first stage of other versions.

⁴ Author's correction from 786 kN to represent mean thrust.

to be 12.6 sq.m for all three stages and the drag coefficient was estimated (eqn 2.70) by assuming a semi-cone angle of 30 degrees. If M_4 denotes the useful mass placed into orbit after dropping the third stage then the payload mass M_0 is given by

$$M_0 = M_4 - M_{veb} \quad (2.83)$$

where M_{veb} is the mass of the vehicle equipment bay (520 kg). The single launch version 40 does not require the "SPELDA". The term payload refers to the spacecraft (including its propulsion system) inserted into GTO; a subsequent manoeuvre at apogee is of course needed to achieve the circular orbit at radius 42164 km.

The initial conditions are as follows:

longitude τ	307.23 deg
latitude δ	5.53 deg
radius r	6378.14 km
air velocity v	0.0 km/s
flight path azimuth from North χ	90.0 deg
flight path elevation γ	90.0 deg

After a simulated vertical ascent for 28 seconds (approximately 1.3 km altitude), the tip-over is implemented in this computation by a small step change θ_t in the flight path elevation angle. This angle and the assumed mass M_4 are inputs to the whole computation which provide the outputs of apogee R_{max} and perigee R_{min} . By running two extra simulations with perturbed values of θ_t and M_4 respectively, the partial derivatives of R_{max} and R_{min} with respect to θ_t and M_4 can be generated. In this way a two-dimensional Newton-Raphson process can be employed to iterate towards the desired values of the apogee and perigee.

The results from such computations are illustrated in Figure 2.4 which shows (against longitude for the equatorial trajectory) (a) altitude with crosses on that plot denoting changes of stage, and (b) the pitch angle θ of the longitudinal axis of the vehicle. The required tip-over was 2.96 degrees

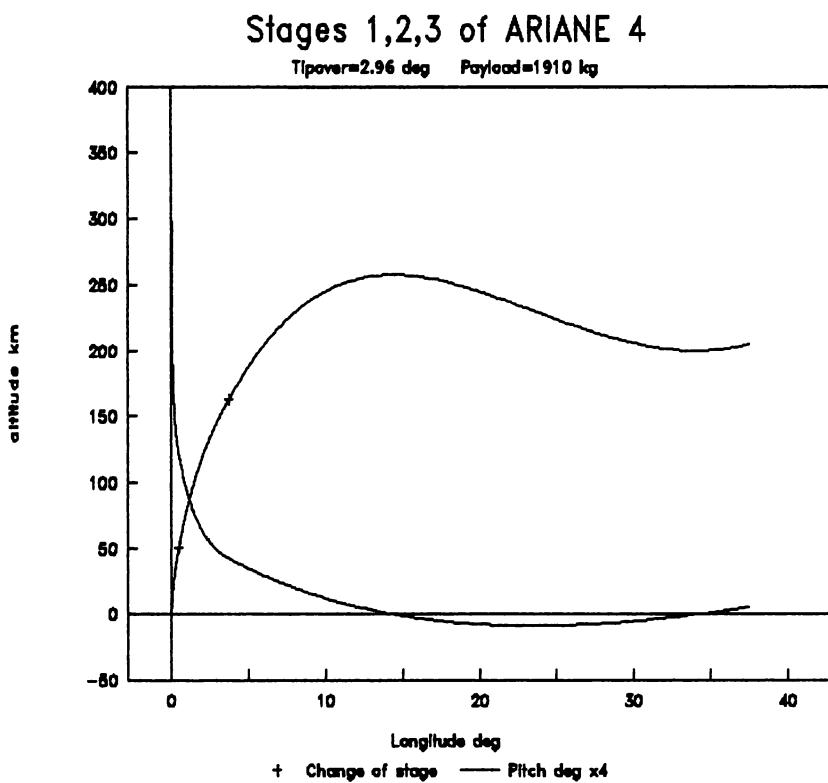


Fig. 2.4. Simulated Launch of the Ariane 40 (Nominal Trajectory)

and the M_4 of 2430 kg, corresponds to a payload of 1910 kg. This number and the illustrated launch trajectory for the three stages are essentially in agreement with the known performance of the Ariane 40.

2.7 Optimization of the Ascent

2.7.1 Solution for the Flat Earth

Consider first a simple case for which the exact optimum trajectory can be calculated, viz. rocket powered ascent with constant thrust acceleration over a flat Earth. The relevant equations of motion can be written down directly as follows where θ is the angle of thrusting above the horizontal.

$$\begin{aligned}\dot{x}_1 &= (F/M) \cos \theta \\ \dot{x}_2 &= (F/M) \sin \theta - g \\ \dot{x}_3 &= x_1 \\ \dot{x}_4 &= x_2\end{aligned}\tag{2.84}$$

x_3 and x_4 are down-range distance and altitude respectively; (F/M) and g are approximated as constants.

Let us seek to maximize the horizontal velocity x_1 at a fixed terminal time t_f subject to the terminal constraints

$$\begin{aligned}x_2 &= x_1 \tan \gamma_f \\ x_4 &= h_f\end{aligned}\tag{2.85}$$

i.e. the final flight path angle γ_f and altitude h_f are specified. Since the final velocity is equal to $x_1(t_f)/\cos \gamma_f$, it follows that this formulation is equivalent to maximizing final velocity for a given mass. It can be shown that this is the same as maximizing the mass for a given final velocity.

This problem is of the type that can be handled by the classical calculus of variations or equivalently by the more modern formulation of optimal control. Readers not familiar with the latter should skip the following manipulations, otherwise the solution is derived by substituting into the standard formulae given in a text such as Bryson and Ho (section 2.5, Ref. [8]). If the equations of the state vector are

$$\dot{x} = f(x, \theta)\tag{2.86}$$

then the equations of the costate vector are

$$\dot{\lambda} = - \left(\frac{\partial f}{\partial \theta} \right)^T \lambda\tag{2.87}$$

Thus the costate equations become

$$\begin{aligned}\dot{\lambda}_1 &= -\lambda_3 \\ \dot{\lambda}_2 &= -\lambda_4 \\ \dot{\lambda}_3 &= 0 \\ \dot{\lambda}_4 &= 0\end{aligned}\tag{2.88}$$

The function to be maximized is

$$\Phi = -x_1(t_f)\tag{2.89}$$

and the constraints to be zeroed are

$$\begin{aligned}\Psi_1 &= x_2(t_f) - x_1(t_f) \tan \gamma_f \\ \Psi_2 &= x_4(t_f) - h_f\end{aligned}\tag{2.90}$$

The terminal values of the costate vector must satisfy

$$\lambda^T(t_f) = \left(\frac{\partial \Phi}{\partial x} + v^T \frac{\partial \Psi}{\partial x} \right)\tag{2.91}$$

where ν is a vector of Lagrangian multipliers. Therefore the final values at t_f of the costates are given by

$$\begin{aligned}\lambda_1 &= -1 - \nu_1 \tan \gamma_f \\ \lambda_2 &= \nu_1 \\ \lambda_3 &= 0 \\ \lambda_4 &= \nu_2\end{aligned}\tag{2.92}$$

This problem is sufficiently simple that the set 2.88 (subject to the final values 2.92) can be integrated to yield the solutions

$$\begin{aligned}\lambda_1 &= -1 - \nu_1 \tan \gamma_f \\ \lambda_2 &= \nu_1 + \nu_2(t_f - t) \\ \lambda_3 &= 0 \\ \lambda_4 &= \nu_2\end{aligned}\tag{2.93}$$

If

$$H = \lambda^T f\tag{2.94}$$

then the optimality condition is

$$\frac{\partial H}{\partial \theta} = 0\tag{2.95}$$

In this case

$$H = \lambda_1(F/M) \cos \theta + \lambda_2[(F/M) \sin \theta - g] + \lambda_3 x_1 + \lambda_4 x_2\tag{2.96}$$

and the optimality condition yields

$$-\lambda_1(F/M) \sin \theta + \lambda_2(F/M) \cos \theta\tag{2.97}$$

or

$$\tan \theta = \lambda_2 / \lambda_1 = \frac{-\nu_1 - \nu_2(t_f - t)}{1 + \nu_1 \tan \gamma_f} \quad (2.98)$$

The constants ν_1 and ν_2 have to be chosen so that the end conditions 2.90 are satisfied. Two other equivalent constants could be used, eg. θ_0 and c in

$$\theta = \arctan(\tan \theta_0 - ct) \quad (2.99)$$

Once the constants θ_0 and c have been adjusted to satisfy the terminal conditions, eqn 2.99 is the optimal solution. It will be referred to as a *linear pitch steering law* in which (for small angles) the arctan function can be ignored if desired. The adjustment of the constants to satisfy the terminal conditions would be by means of a two-dimensional Newton-Raphson process employing partial derivatives of the end conditions 2.85 with respect to θ_0 and c . In this case analytical expressions are possible but generally perturbed numerical integrations of the state equations (including eqn 2.99) would be necessary.

2.7.2 Improvement of the Nominal Trajectory

Numerical methods based on optimal control theory and constrained optimization algorithms have been applied to compute ascent trajectories but in practice simpler calculations are usually sufficiently accurate, bearing in mind that some rocket parameters such as thrust level and aerodynamic coefficients are not known accurately beforehand. Some form of in-flight correction is necessary and therefore it is unrealistic to attempt very precise optimization in advance.

With this in mind, the computations of section 2.6 (Mode 1) have been repeated as follows (Mode 2).

1. Adjust mass M_4 and tip-off angle θ_t as before to satisfy the two desired end conditions (apogee and perigee).
2. After the first stage, if θ_0 is the pitch angle at the end of that stage, the pitch angle is subsequently adjusted according to a linear pitch steering law, viz.

$$\theta = (1 - \mu)\theta_0 + \mu\theta_1 \quad ; \quad \mu = (t - t_1)/(t_f - t_1) \quad (2.100)$$

where t_1 is the time at the end of the first stage.

Thus in the first stage as before the aerodynamic angle of attack is essentially zero. *This is virtually dictated in practice* in order to limit aerodynamic drag and lift, which would result in unacceptably high lateral accelerations. After the first stage, when the vehicle is largely out of the atmosphere, the pitch angle is reduced as a linear function of time to a final set value θ_1 , ie. the approximation to the optimal solution of section 2.7.1 is inserted.

When the parameter θ_1 is adjusted by trial-and-error to maximize the payload the following results are obtained, where the previous results of section 2.6 (Mode 1) are repeated as a matter of convenience. Mode 2 refers to this revised computation.

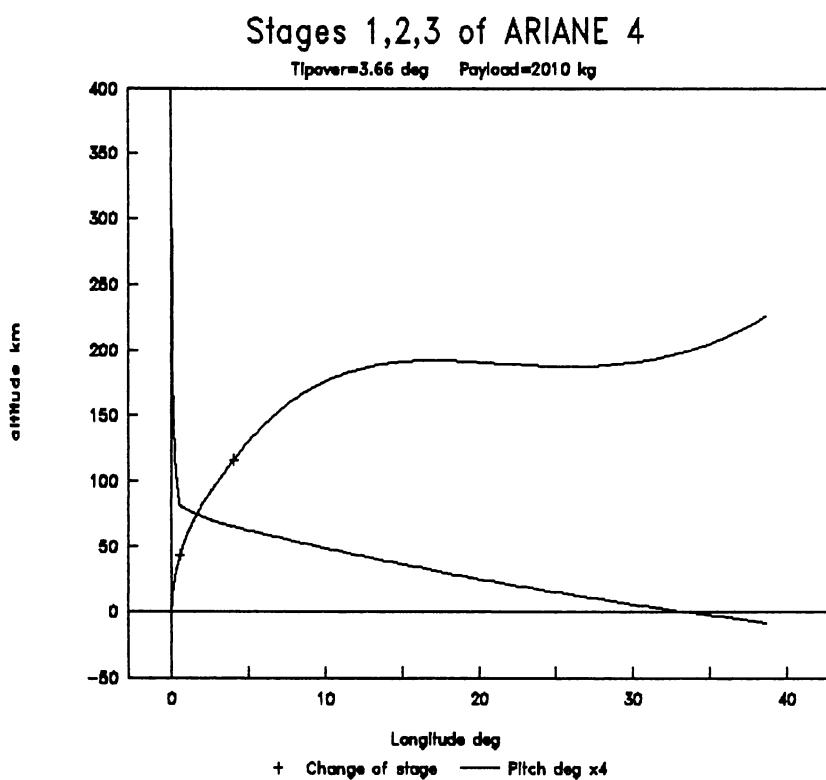


Fig. 2.5. Simulated Launch of the Ariane 40 (Improved Trajectory)

Mode 1 : $\theta_t=2.959$ deg, $M_4=2430.1$ kg, payload=1910 kg.

Mode 2 : $\theta_t=3.663$ deg, $\theta_1=-2.0$ deg, $M_4=2530.2$ kg, payload=2010 kg.

Thus a 1 per cent increase in the payload has been achieved as a result of the steering law derived for the flat Earth. The ascent trajectories and profile of the pitch angle is shown in Figure 2.5 for Mode 2; see Figure 2.4 for Mode 1.

2.8 In-flight Guidance

The small improvement with the linear pitch steering law was apparently beneficial and permitted a closer approach to the true optimal ascent trajectory, but (more important) it was introduced because it has been the basis for the kind of on-line guidance commonly adopted in practice, eg. Ariane (Ref [55]), Saturn (Ref [24]), and Atlas/Centaur (Ref [7]). As mentioned above, an actual ascent will differ from a pre-computed nominal trajectory because of winds, variations in propulsive performance, uncertainties in aerodynamic coefficients, etc. If such effects were uncorrected then the dispersion in the position and velocity vector at final-stage termination would usually be unacceptable. In-flight or *on-line guidance* is therefore required and is typically realised as follows.

1. The first stage uses only open-loop guidance to burn-out, ie. the longitudinal axis of the rocket follows a pre-computed schedule in pitch and yaw maintaining approximately zero aerodynamic angle of attack.
2. The second and third stages have closed-loop guidance in which parameters of a pitch (and possibly yaw) steering law are adjusted, together with the time of final stage cut-off, to achieve the desired final orbit despite in-flight variations.

Such guidance has been included as an example in the simulations of Ariane 40 as described above. Specifically Mode 3 is defined as the above *on-line guidance*. The pitch schedule in the first stage is that generated in the mode 2 simulation. The apogee and perigee of the final orbit are obtained by adjustments of θ_0 (eqn 2.100) and the time of third stage cut-off. The parameter θ_1 (eqn 2.100) is fixed at the optimum value of -2.0 degrees. The first simulation (Mode 3a) represents a revised nominal trajectory with a slightly shorter time of third stage burning, viz. 579.5 seconds compared to the previous maximum of 584.0 seconds. This allows a capability for adjusting final stage cut-off to correct in-flight variations. The second simulation (Mode 3b) illustrates how a +1 per cent (above nominal) thrust in stage 1 can be corrected by adjusting θ_0 and the time of third stage cut-off.

Mode 3a : $\theta_0=20.149$ deg, stage 3 cut-off=579.50 seconds, $M_4=2450.0$ kg, payload=1910 kg.

Mode 3b : $\theta_0=19.320$ deg, stage 3 cut-off=576.75 seconds, $M_4=2450.0$ kg, payload=1910 kg.

2.9 The Guidance and Navigation System

An extensive treatment of the practical aspects of in-flight guidance and navigation are beyond the scope of this text but the principal features of such systems can be summarized. First more detail is added to the guidance simulated in section 2.8.

Let R_1, R_2 represent any pair of orbital parameters to be achieved at a final time t_f . However choose t_f to ensure that it is always after final (third) stage cut-off at time T_3 , allowing for variations in the latter. If \mathbf{x} denotes the vector of position and velocity and Δ signifies perturbations

$$\begin{bmatrix} \Delta R_1 \\ \Delta R_2 \end{bmatrix} = A(t_f, t) \Delta \mathbf{x}(t) \quad (2.101)$$

where the matrix A can be computed from perturbed trajectories. Similarly a matrix B can be generated to relate perturbations in the guidance parameters θ_0 (eqn 2.100) and cut-off time T_3 to ΔR_1 and ΔR_2

$$\begin{bmatrix} \Delta R_1 \\ \Delta R_2 \end{bmatrix} = B(t_f, t) \begin{bmatrix} \Delta \theta_0 \\ \Delta T_3 \end{bmatrix} \quad (2.102)$$

Combine the two equations to give

$$\begin{bmatrix} \Delta \theta_0 \\ \Delta T_3 \end{bmatrix} = D(t_f, t) \Delta \mathbf{x}(t) \quad (2.103)$$

where

$$D(t_f, t) = B^{-1}(t_f, t) A(t_f, t) \quad (2.104)$$

In the simplest situation, the (2×2) matrix D can be pre-computed along the nominal trajectory and stored in the onboard computer at frequent intervals of time or possibly down-range distance. Given a determination of the current position and velocity vector (compared to the nominal trajectory), eqn 2.103 can be employed to calculate the required adjustments to the two guidance parameters θ_0 and T_3 . The above analysis can be extended to adjust also θ_1 , in order to include a third orbital variable to be adjusted. If the departures from the nominal trajectory are too great then eqn 2.103 must be applied iteratively, in which revised trajectories and sensitivity matrices are computed in-flight. In this way the required pitch and yaw angles (and the expected cut-off time) are updated at frequent intervals in flight. The desired pitch and yaw angles become reference signals to the attitude control system which is driven in the rocket-powered phase by thrust vector control of the rocket engine. This is *the guidance system*.

The above scheme applied to guidance parameters derived from the flat-Earth optimization is *essentially the basis for on-line guidance (after the atmospheric phase) applied to most launch vehicles to date*, eg. Saturn (Ref. [24]), Ariane (Ref. [55]), Atlas/Centaur, and Titan/Centaur (Ref. [7]). There

is scope for parameterization of the computations to reduce the in-flight computing load, eg. Ariane stores onboard parameters as a function of potential and kinetic energy (Ref. [55]). Simpler equations of motion were employed in the onboard Saturn guidance computer (Ref. [24]), but such simplifications are no longer necessary with modern processors, eg. the guidance computer of Titan/Centaur (Ref. [7]) performs real-time integration of the dynamical equations.

The guidance system requires at regular intervals a determination of the current position and velocity vector. This is the job of *the navigation system*. It would be possible to determine the position and velocity of a vehicle during the launch phase either by radar measurements of range and range-rate (doppler) from several ground tracking sites, or by the Global Positioning System GPS (appendix A.6) which relies on a network of high altitude satellites. This has not been the primary system to date; instead the reliability of autonomous inertial navigation systems has been preferred.

An inertial navigation system depends on a set of mutually perpendicular accelerometers and a set of mutually perpendicular gyros. Accelerometers will measure nothing inside a spacecraft in a free non-propulsive orbit outside the atmosphere, but in the launch phase they can sense accelerations due to propulsive and aerodynamic forces. An accelerometer is typically a small pendulum with one degree of freedom. Movement along the sensitive direction is restrained electromagnetically such that the feedback electrical current is proportional to the acceleration. Position and velocity are obtained by integrating the accelerometer outputs and including the calculated effects of the gravitational field. Three single-degree-of-freedom gyros mounted about mutually perpendicular axes provide rates of turn which are integrated to yield the angles which define the orientation of the axes in inertial space. Both accelerometers and gyros are commonly mounted on gas bearings.

All the early inertial navigation systems mounted the accelerometers and gyros on a stabilized three-gimbal platform, which maintained its orientation (by means of the gyro outputs) in inertial space despite the angular motion of the vehicle carrying the platform. Small drift errors accumulate gradually which must be corrected in aircraft systems, but they can be tolerated for the relatively short duration of a rocket launch phase. Since the axes of the accelerometers are invariant with respect to inertial axes, frequent computation of sines and cosines is not necessary to carry out orthogonal (rotational) transformations, an important point with the relatively primitive computers originally available for flight applications. Even though inertial platforms have not disappeared, the trend is however towards *strapdown inertial systems* where both the accelerometers and gyros are fixed in the vehicle. Transformations are required at frequent intervals to transform from the moving vehicle axes to inertial axes but this is well within the scope of modern computers. The reader should nonetheless be aware that processing power for space applications lags significantly that for ground use, due to the delay of

space qualification and the effects of radiation in space. Ring Laser gyros are also a feature of new systems. As a result of eliminating the heavy platforms of earlier navigation systems and the reduced size of computers, a modern inertial navigation unit is lighter and consumes less power than earlier designs, eg. for the Titan/Centaur (Ref. [60]) the mass and power of the navigation unit has been reduced from 68 kg (335 watts) to 28 kg (120 watts).

2.10 Software

The simulations of the launch phase were carried out by means of C++ programs grouped together in the file ASCENT.CPP which has been deposited in the Internet FTP server of the publisher as specified in the preface. The routines included in that file are as follows.

Main : (section 2.4.1) main program to compute launch trajectories to satisfy two end conditions R_1, R_2 . Input vector of parameters is \mathbf{x} .

launch : (section 2.4.1) given \mathbf{x} generate R_1, R_2 .

constraints : (section 2.4.1) given \mathbf{x} generate errors in R_1, R_2 .

gradients : (section 2.4.1) generate partial derivatives of errors in R_1, R_2 with respect to \mathbf{x} .

Runge_Kutta4 : (section 2.4.1) Fourth order Runge Kutta numerical integration.

RHS : (section 2.4.1) formation of the right hand sides of the six first order differential equations.

atmosphere : (section 2.4.4) US Standard Atmosphere 1962.

translate : (section 2.4.1) copying components of the state vector into variables used in RHS etc., or vice versa.

Earth_radius : (section 2.4.2) Earth radius as a function of latitude.

plot_xy : screen plotting module.

gravity : (section 2.4.2) gravitational force as a function of latitude and radius.

drag : (section 2.4.4) aerodynamic force as a function of the position and velocity vectors.

The file MATRIX.CPP of matrix routines is also required.

CHAPTER 3

EARTH SATELLITES: PERTURBATIONS AND MANOEUVRES

3.1 Introduction

After a spacecraft has been placed in an operational orbit about the Earth, subsequent manoeuvres will be required, either to correct the orbit or as demanded by the mission, eg. in order to rendezvous with another vehicle. Correction of orbits becomes necessary as a result of perturbations from the non-spherical shape of the Earth, the influence of the Sun, the Moon and any residual atmosphere, although the latter is significant only at relatively low altitudes (below 100 km). This chapter includes an analysis of perturbations and the necessary corrective manoeuvres but geostationary orbits are treated separately because of their singular nature and tight tolerances.

Satellite rendezvous also calls for a completely different kind of analysis.

3.2 Variation of the Orbital Elements

If a satellite is influenced only by the gravitational field of a point mass or a spherical body, then it was shown in Chapter 1 that the orbit is an ellipse, parabola or hyperbola. Given the position and velocity vectors at any time, the orbital elements can be calculated (section 1.3.2) but these elements will be invariant. However, if the motion of the satellite is perturbed in any way by other forces, then the calculation of the orbital elements will yield a different set of values over an interval of time. This orbit with varying parameters is called an *osculating orbit*. Instead of computing the perturbed orbit from the equations of motion in position and velocity, it is possible to treat the orbital elements as the dependent variables of a set of first order differential equations. At any time, the position and velocity vectors can be calculated from the set of evolving orbital elements (section 1.3.1). These equations which describe variation of the orbital elements are now derived.

3.2.1 Rates of Change of the Elements

In the following analysis let δ indicate a change in an orbital variable due to the application of a vector f of accelerations *other than due to the spherically*

symmetrical central gravitational field. The notation of chapter 1 is employed and we start by noting, from eqn 1.23, the change in the energy per unit mass over a time interval δt .

$$\delta E = \frac{\mu \delta a}{2a^2} = \mathbf{v} \cdot \mathbf{f} \delta t \quad (3.1)$$

In the limit this gives the rate of change of energy as

$$\frac{da}{dt} = \frac{2a^2}{\mu} (\mathbf{v} \cdot \mathbf{f}) \quad (3.2)$$

Since from eqn 1.33

$$h^2 = (\mathbf{r} \times \mathbf{v}) \cdot (\mathbf{r} \times \mathbf{v}) \quad (3.3)$$

the rate of change of angular momentum can be similarly calculated.

$$\begin{aligned} \frac{dh}{dt} &= (\mathbf{r} \times \mathbf{v}) \cdot (\mathbf{r} \times \mathbf{f}) / h \\ &= \mathbf{r} \times (\mathbf{f} \times \mathbf{r}) \cdot \mathbf{v} / h \\ &= [r^2(\mathbf{v} \cdot \mathbf{f}) - (\mathbf{r} \cdot \mathbf{v})(\mathbf{r} \cdot \mathbf{f})] / h \end{aligned} \quad (3.4)$$

having applied the standard identity for a vector triple product.

In order to obtain the rate of change of eccentricity, combine eqns 1.15 and 1.22

$$h^2 = \mu a(1 - e^2) \quad (3.5)$$

Differentiate using eqns 3.2 and 3.4 to yield

$$\frac{de}{dt} = \frac{1}{\mu a e} [(pa - r^2)(\mathbf{v} \cdot \mathbf{f}) + (\mathbf{r} \cdot \mathbf{v})(\mathbf{r} \cdot \mathbf{f})] \quad (3.6)$$

To calculate the out-of-plane elements note from eqn 1.29

$$\mathbf{h} = \mathbf{r} \times \mathbf{v} = h \begin{bmatrix} \sin i \sin \Omega \\ -\sin i \cos \Omega \\ \cos i \end{bmatrix} \quad (3.7)$$

Differentiate this equation and arrange the result in the form

$$\mathbf{r} \times \mathbf{f} = \Gamma_2 \begin{bmatrix} h \sin i \dot{\Omega} \\ -h \dot{i} \\ h \end{bmatrix} \quad (3.8)$$

where

$$\Gamma_2 = \begin{bmatrix} \cos \Omega & -\sin \Omega \cos i & \sin \Omega \sin i \\ \sin \Omega & \cos \Omega \cos i & -\cos \Omega \sin i \\ 0 & \sin i & \cos i \end{bmatrix} \quad (3.9)$$

which is an orthogonal matrix; transpose equals the inverse. Therefore

$$\begin{bmatrix} h \sin i d\Omega/dt \\ -hdi/dt \\ dh/dt \end{bmatrix} = \Gamma_2^T (\mathbf{r} \times \mathbf{f}) \quad (3.10)$$

but note that the elements Ω and i refer to the same axes as the vectors on the right hand side of the above equation.

It remains to derive the rate of change of the true anomaly θ but, in order to distinguish the change due only to \mathbf{f} , we employ the symbol $\delta\tilde{\theta}$ and similarly in writing down the limiting rate of change. Thus by applying perturbations to eqn 1.14

$$r\delta e \cos \theta - re \sin \theta \delta\tilde{\theta} = 2h\delta h/\mu \quad (3.11)$$

or in the limit

$$re \sin \theta \frac{d\tilde{\theta}}{dt} = r \cos \theta \frac{de}{dt} - \frac{2h}{\mu} \frac{dh}{dt} \quad (3.12)$$

Since e and h have already been evaluated, the above suffices to define the rate of change of $\tilde{\theta}$. There are several equivalent forms but it is perhaps preferable to complete this calculation as follows. From eqns 1.15 and 1.25

$$h(\mathbf{r}, \mathbf{v}) = r\mu e \sin \theta \quad (3.13)$$

Differentiate to yield

$$re \cos \theta \frac{d\tilde{\theta}}{dt} = -r \sin \theta \frac{de}{dt} + \frac{p}{h} \left[(\mathbf{r} \cdot \mathbf{f}) + (\mathbf{r} \cdot \mathbf{v}) \frac{1}{h} \frac{dh}{dt} \right] \quad (3.14)$$

Combine eqns 3.12 and 3.14

$$re \frac{d\tilde{\theta}}{dt} = (p/h) \cos \theta (\mathbf{r} \cdot \mathbf{f}) + \left[\frac{p}{h^2} \cos \theta (\mathbf{r} \cdot \mathbf{v}) - \frac{2h}{\mu} \sin \theta \right] \frac{dh}{dt} \quad (3.15)$$

$$\begin{aligned} \left[\frac{p}{h^2} \cos \theta (\mathbf{r} \cdot \mathbf{v}) - \frac{2h}{\mu} \sin \theta \right] &= -\sin \theta [p + r(1 + e \cos \theta) - er \cos \theta] / h \\ &= -\sin \theta (p + r) / h \end{aligned}$$

or finally

$$\frac{d\tilde{\theta}}{dt} = \frac{1}{reh} \left[p \cos \theta (\mathbf{r} \cdot \mathbf{f}) - (p + r) \sin \theta \frac{dh}{dt} \right] \quad (3.16)$$

The variation of the argument of pericentre is obtained by starting with eqns 1.28 and 1.41, viz.

$$\mathbf{i}_x \cdot \mathbf{r} = [\cos \Omega \quad \sin \Omega \quad 0] \cdot \mathbf{r} = r \cos(\theta + \omega) \quad (3.17)$$

Differentiation results in

$$[-\sin \Omega \quad \cos \Omega \quad 0] \cdot \mathbf{r} \frac{d\Omega}{dt} = -r \sin(\theta + \omega) \left(\frac{d\tilde{\theta}}{dt} + \frac{d\omega}{dt} \right) \quad (3.18)$$

From eqns 1.26, 1.27, 1.30, and 1.31 we can verify that

$$\begin{aligned} \begin{bmatrix} -\sin\Omega & \cos\Omega & 0 \end{bmatrix} \cdot \mathbf{r} &= \begin{bmatrix} -\sin\Omega & \cos\Omega & 0 \end{bmatrix} \Gamma_2 \Gamma_1 \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix} \\ &= \cos i \sin(\theta + \omega) r \end{aligned}$$

Therefore from eqn 3.18

$$\cos i \frac{d\Omega}{dt} = - \left(\frac{d\tilde{\theta}}{dt} + \frac{d\omega}{dt} \right) \quad (3.19)$$

or

$$\frac{d\omega}{dt} = -\cos i \frac{d\Omega}{dt} - \frac{d\tilde{\theta}}{dt} \quad (3.20)$$

We now have the required six equations (3.2, 3.6, 3.10 for Ω and i , 3.16, and 3.20) for variations of the orbital elements; eqn 3.4 for h is an auxiliary equation. The final form of the equations depends on the axes in use.

3.2.2 Use of Radial and Transverse Components

Let us first employ the orbit plane axes ($\mathbf{i}_r, \mathbf{i}_\theta, \mathbf{i}_z$) of Figure 1.1 and section 1.3.1. Then

$$\mathbf{r} \cdot \mathbf{f} = r f_r \quad (3.21)$$

and

$$\mathbf{r} \cdot \mathbf{v} = r v \sin \gamma \quad (3.22)$$

where γ is the angle between the velocity vector \mathbf{v} and \mathbf{i}_θ measured clockwise from the latter, ie. it is a negative rotation about \mathbf{i}_z . It follows from eqn 1.25 that

$$\tan \gamma = \frac{e \sin \theta}{1 + e \cos \theta} \quad (3.23)$$

Note also from eqn 1.25

$$\begin{aligned} \mathbf{v} \cdot \mathbf{f} &= \sqrt{\mu/p} [e f_r \sin \theta + (1 + e \cos \theta) f_\theta] \\ &= \frac{\mu e}{h} \sin \theta f_r + \frac{h}{r} f_\theta \end{aligned} \quad (3.24)$$

Equation 3.10 has the vector $(\mathbf{r} \times \mathbf{f})$ which must be in the ijk axes of Figure 1.2, whereas we are here using components in the $r\theta z$ axes of that figure. The required transformation is at eqn 1.31, hence eqn 3.10 can be rewritten as

$$\begin{aligned}
 \begin{bmatrix} h \sin id\Omega/dt \\ -hdi/dt \\ dh/dt \end{bmatrix} &= \Gamma_2^T (\mathbf{r} \times \mathbf{f})_{ijk} \\
 &= \Gamma_2^T \Gamma_2 \Gamma_1 (\mathbf{r} \times \mathbf{f})_{r\theta z} \\
 &= \Gamma_1 \begin{bmatrix} 0 \\ -rf_z \\ rf_\theta \end{bmatrix} \tag{3.25}
 \end{aligned}$$

3.2.3 Use of Tangential and Normal Components

First note from eqn 1.25 that

$$v = \sqrt{\frac{\mu}{p}}(1 + 2e \cos \theta + e^2) \tag{3.26}$$

and applying also eqn 3.23

$$\begin{aligned}
 v \sin \gamma &= \sqrt{\frac{\mu}{p}} e \sin \theta \\
 v \cos \gamma &= \sqrt{\frac{\mu}{p}} (1 + e \cos \theta) \tag{3.27}
 \end{aligned}$$

The required scalar products are then obtained as

$$\begin{aligned}
 \mathbf{r} \cdot \mathbf{f} &= r(f_t \sin \gamma - f_n \cos \gamma) \\
 &= \frac{\mu re}{vh} \sin \theta f_t - \frac{h}{v} f_n \tag{3.28}
 \end{aligned}$$

$$\mathbf{v} \cdot \mathbf{f} = vf_t \tag{3.29}$$

The result of eqn 3.25 is modified by means of the transformation from tnz to $r\theta z$ axes, viz.

$$\Gamma_3 = \begin{bmatrix} \sin \gamma & -\cos \gamma & 0 \\ \cos \gamma & \sin \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.30}$$

Since

$$(\mathbf{r} \times \mathbf{f})_{tnz} = \begin{bmatrix} -rf_z \cos \gamma \\ -rf_z \sin \gamma \\ rf_n \sin \gamma + rf_t \cos \gamma \end{bmatrix} \tag{3.31}$$

eqn 3.25 becomes for the tnz axes

$$\begin{bmatrix} h \sin id\Omega/dt \\ -hdi/dt \\ dh/dt \end{bmatrix} = \Gamma_1 r \begin{bmatrix} 0 \\ -f_z \\ f_n \sin \gamma + f_t \cos \gamma \end{bmatrix} \tag{3.32}$$

3.2.4 Summary of Equations in Radial-Transverse Axes

$$\frac{da}{dt} = 2a^2 \left(e \sin \theta f_r + \frac{p}{r} f_\theta \right) / h \quad (3.33)$$

$$\frac{de}{dt} = \frac{p}{h} \left[\sin \theta f_r + \left(\cos \theta + \frac{r}{p} (\cos \theta + e) \right) f_\theta \right] \quad (3.34)$$

$$\frac{dh}{dt} = r f_\theta \quad (3.35)$$

$$\frac{d\tilde{\theta}}{dt} = \frac{1}{eh} [p \cos \theta f_r - (p + r) \sin \theta f_\theta] \quad (3.36)$$

$$\frac{d\Omega}{dt} = \frac{r \sin(\omega + \theta)}{h \sin i} f_z \quad (3.37)$$

$$\frac{di}{dt} = \frac{r \cos(\omega + \theta)}{h} f_z \quad (3.38)$$

$$\frac{d\omega}{dt} = -\cos i \frac{d\Omega}{dt} - \frac{d\tilde{\theta}}{dt} \quad (3.39)$$

3.2.5 Summary of Equations in Tangential-Normal Axes

$$\frac{da}{dt} = \frac{2a^2 v}{\mu} f_t \quad (3.40)$$

$$\frac{de}{dt} = \frac{1}{v} \left[2(e + \cos \theta) f_t - \frac{r}{a} \sin \theta f_n \right] \quad (3.41)$$

$$\frac{dh}{dt} = \frac{rh}{pv} \left[\frac{p}{r} f_t + e \sin \theta f_n \right] \quad (3.42)$$

$$\frac{d\tilde{\theta}}{dt} = -\frac{1}{ev} \left[2 \sin \theta f_t + \left(2e + \frac{r}{a} \cos \theta \right) f_n \right] \quad (3.43)$$

Equations 3.37, 3.38, and 3.39 are unchanged.

3.2.6 Equations Applicable at Zero Eccentricity

At zero eccentricity ω is not defined and the above equations are singular. Since near-circular orbits are so common, let us note the resulting equations when substitutions are made to avoid the singularity. Eliminate e and ω by means of

$$e_x = e \cos \omega \quad e_y = e \sin \omega \quad (3.44)$$

$$\theta^* = \theta + \omega \quad (3.45)$$

It follows that

$$p = a(1 - e_x^2 - e_y^2) \quad (3.46)$$

$$e \cos \theta = e \cos(\theta^* - \omega) = e_x \cos \theta^* + e_y \sin \theta^* \quad (3.47)$$

Therefore

$$r = \frac{a(1 - e_x^2 - e_y^2)}{1 + e_x \cos \theta^* + e_y \sin \theta^*} \quad (3.48)$$

and

$$v = \frac{\mu}{h} \sqrt{1 + e_x^2 + e_y^2 + 2e_x \cos \theta^* + 2e_y \sin \theta^*} \quad (3.49)$$

Using the equations of section 3.2.5 and eqn 3.20, evaluate the derivatives of e_x and e_y to yield

$$\frac{de_x}{dt} = \frac{2}{v}(e_x + \cos \theta^*) f_t - \frac{1}{v} \left(2e_y + \frac{r}{a} \sin \theta^* \right) f_n + \frac{re_y \sin \theta^*}{h \tan i} f_z \quad (3.50)$$

$$\frac{de_y}{dt} = \frac{2}{v}(e_y + \sin \theta^*) f_t + \frac{1}{v} \left(2e_x + \frac{r}{a} \cos \theta^* \right) f_n - \frac{re_x \sin \theta^*}{h \tan i} f_z \quad (3.51)$$

The rates of change of a , i and Ω are unchanged but they are repeated for completeness.

$$\frac{da}{dt} = \frac{2a^2 v}{\mu} f_t \quad (3.52)$$

$$\frac{di}{dt} = \frac{r \cos \theta^*}{h} f_z \quad (3.53)$$

$$\frac{d\Omega}{dt} = \frac{r \sin \theta^*}{h \sin i} f_z \quad (3.54)$$

Finally, with the help of eqn 3.39 calculate the *total rate of change* of θ^* as

$$\frac{d\theta^*}{dt} = \frac{d\theta}{dt} + \frac{d\tilde{\theta}}{dt} + \frac{d\omega}{dt} = h/r^2 - \cos i \frac{d\Omega}{dt} \quad (3.55)$$

If it is desired to use the above equations in conjunction with the acceleration vector in equatorial ijk axes, note

$$\mathbf{f}_{tnz} = \Gamma_3^T \Gamma_1^T \Gamma_2^T \mathbf{f}_{ijk} \quad (3.56)$$

3.2.7 The Singularity at Zero Inclination

At zero inclination the longitude of the ascending node Ω is not defined. There are essentially two methods of dealing with this singularity as follows.

1. employ a completely different set of orbital elements (Ref. [3]) or,
2. use a reference frame rotated through 90 degrees so that zero inclination becomes 90 degrees.

The author has preferred the latter; there are three slightly different forms of pre-processing as specified below. First define an orthogonal transformation

$$U = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad (3.57)$$

Any vector \mathbf{v} is transformed into \mathbf{v}' by

$$\mathbf{v}' = U\mathbf{v} \quad (3.58)$$

1. If the starting conditions are position and vector simply transform to the new reference frame by means of U .
2. If the starting conditions are orbital elements, calculate position and velocity (despite zero inclination) by means of the algorithm specified in section 1.3.1. Transform position and velocity to the new reference frame by means of U and then compute the new orbital elements using the algorithm delineated in section 1.3.2. The latter will be valid because the new inclination will be 90 degrees.
3. If the starting conditions are orbital elements, they can alternatively be transformed as follows without resorting to position and velocity as follows.
 - a) Calculate the unit vectors $(\mathbf{i}_x, \mathbf{i}_y, \mathbf{i}_z)$ of Figure 1.2 from (i, Ω) ; see section 1.3.1.
 - b) Form \mathbf{u} as the unit eccentricity vector from (e_x, e_y) or ω by

$$\mathbf{u} = \cos \omega \mathbf{i}_x + \sin \omega \mathbf{i}_y \quad (3.59)$$
 Set ω to zero if the eccentricity is zero.
 - c) Using U transform to the new reference frame to obtain \mathbf{i}'_z and \mathbf{u}' .
 - d) The new (i', Ω') are obtained from \mathbf{i}'_z (section 1.3.2) which in turn yields \mathbf{i}'_x .
 - e) The new ω' is calculated from

$$\sin \omega' = \mathbf{i}'_z \cdot (\mathbf{i}'_x \times \mathbf{u}') \quad (3.60)$$

$$\cos \omega' = i'_x \cdot \mathbf{u}' \quad (3.61)$$
 - f) If using θ^*

$$\theta^*(new) = \theta^*(old) + (\omega' - \omega) \quad (3.62)$$

3.3 Gravitational Field of the Earth

3.3.1 Expansion of Spherical Harmonics

In order to evaluate the acceleration vector due to a non-spherical Earth we need the generalised expansion of the potential function. This is a well known formula and is quoted without derivation; see for example Battin (Ref. [3]) or Kaplan (Ref. [29]). Assume XYZ axes fixed in the Earth (origin at the centre

of gravity) with the X axis at zero longitude, and corresponding conventional spherical polar coordinates except that ψ (not θ) is the angle from the Z axis to the radius vector. The potential function is

$$V(r, \psi, \phi) = \frac{\mu}{r} \left[1 - \sum_{k=2}^{\infty} (R_0/r)^k J_k P_k(\cos \psi) \right] + \frac{\mu}{r} \left[\sum_{k=2}^{\infty} \sum_{j=1}^k (R_0/r)^k P_k^j(\cos \psi) (C_k^j \cos j\phi + S_k^j \sin j\phi) \right] \quad (3.63)$$

where $P_k^j(x)$ is a Legendre function defined in terms of Legendre polynomials by means of

$$P_k^j(x) = (1-x^2)^{1/2} \frac{d^i}{dx^j} P_k(x) \quad (3.64)$$

Note that

$$P_0(x) = 1 \quad P_1(x) = x \quad P_2(x) = (3x^2 - 1)/2 \quad (3.65)$$

The vector of acceleration due to gravity is

$$\mathbf{g}_{r\psi\phi} = \frac{\partial V}{\partial r} \mathbf{i}_r + \frac{1}{r} \frac{\partial V}{\partial \psi} \mathbf{i}_\psi + \frac{1}{r \sin \psi} \frac{\partial V}{\partial \phi} \mathbf{i}_\phi \quad (3.66)$$

The above vector will be required with respect to the XYZ axes fixed in the (rotating) Earth, the transformation being established in the usual way by writing down unit vectors along δr , $\delta\psi$, and $\delta\phi$ in XYZ axes. Thus if (l, m, n) are the direction cosines of the vector \mathbf{r}

$$\mathbf{g}_{XYZ} = \begin{bmatrix} l & n \cos \phi & -\sin \phi \\ m & n \sin \phi & \cos \phi \\ n & (-m \sin \phi - l \cos \phi) & 0 \end{bmatrix} \mathbf{g}_{r\psi\phi} \quad (3.67)$$

where

$$l = \sin \psi \cos \phi \quad m = \sin \psi \sin \phi \quad n = \cos \phi \quad (3.68)$$

3.3.2 Evaluation for an Oblate Earth

Let us first note the coefficients of the first few terms of the expansion 3.63.

$$\begin{aligned} J_2 &= 1082.7 \times 10^{-6} & C_2^1 &= 0 & S_2^1 &= 0 \\ J_3 &= -2.56 \times 10^{-6} & C_2^2 &= 1.57 \times 10^{-6} & S_2^2 &= -0.897 \times 10^{-6} \\ J_4 &= -1.58 \times 10^{-6} & C_3^1 &= 2.10 \times 10^{-6} & S_3^1 &= 0.16 \times 10^{-6} \end{aligned}$$

It is clear that the single dominant term is due to J_2 which arises from an oblate Earth almost symmetrical about the polar Z axis. Therefore for most purposes it is sufficient to employ an expansion only to second order and approximate the Earth to be symmetrical about the Z axis. In this case the simplified expressions become, *ignoring the principal gravity term*,

$$\frac{\partial V}{\partial r} = \frac{3\mu}{r^2} (R_0/r)^2 J_2 (3 \cos^2 \psi - 1)/2 \quad (3.69)$$

$$\frac{1}{r} \frac{\partial V}{\partial \psi} = \frac{3\mu J_2 \sin \psi \cos \psi}{r^2} (R_0/r)^2 \quad (3.70)$$

$$\frac{1}{r \sin \psi} \frac{\partial V}{\partial \phi} = 0 \quad (3.71)$$

Note however that it will be necessary to transform from \mathbf{g} in rotating XYZ axes (fixed in the Earth) to inertial (non-rotating) equatorial axes by means of

$$\mathbf{g}_{XYZ}(\text{fixed}) = \begin{bmatrix} \cos \eta & -\sin \eta & 0 \\ \sin \eta & \cos \eta & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{g}_{XYZ}(\text{rotating}) \quad (3.72)$$

where

$$\eta = \omega_e(t - t_{ref}) \quad (3.73)$$

ω_e being the sidereal rate of rotation of the Earth.

3.3.3 Analytical Results for an Oblate Earth

The calculation of the effects of perturbations is made tedious as a result of transformations needed between different frames of axes. Illustrative results are consequently generated below by means of computer simulations. Analytical results for the effects of Earth oblateness are however possible.

From eqns 3.69 and 3.70 the perturbation vector of accelerations is

$$\mathbf{f} = \frac{3\mu}{2r^2} (R_0/r)^2 J_2 [(3 \cos^2 \psi - 1)\mathbf{i}_r + 2 \sin \psi \cos \psi \mathbf{i}_\psi] \quad (3.74)$$

Instead of using the unit vector \mathbf{i}_ψ substitute the unit vector \mathbf{k} along the equatorial Z axis.

$$\mathbf{i}_\psi = (\cos \psi \mathbf{i}_r - \mathbf{k}) / \sin \psi \quad (3.75)$$

whereupon

$$\mathbf{f} = \frac{3\mu}{2r^2} (R_0/r)^2 J_2 [(5 \cos^2 \psi - 1)\mathbf{i}_r - 2 \cos \psi \mathbf{k}] \quad (3.76)$$

We wish to obtain \mathbf{f} in the $(\mathbf{i}_r, \mathbf{i}_\theta, \mathbf{i}_z)$ axes of Figures 1.1 and 1.2 so \mathbf{k} must be transformed to \mathbf{k}' by means of the transformations of that section.

$$\mathbf{k}' = \Gamma_1^T \Gamma_2^T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sin \theta^* \sin i \\ \cos \theta^* \sin i \\ \cos i \end{bmatrix} \quad (3.77)$$

where θ^* is defined at eqn 3.45. Given that

$$\mathbf{k}' \cdot \mathbf{i}_r = \cos \psi = \sin \theta^* \sin i \quad (3.78)$$

we obtain the perturbation vector in the required axes as

$$\begin{bmatrix} f_r \\ f_\theta \\ f_z \end{bmatrix} = -\frac{3\mu J_2}{2r^2} (R_0/r)^2 \begin{bmatrix} 1 - 3\sin^2 \theta^* \sin i \\ \sin 2\theta^* \sin^2 i \\ \sin \theta^* \sin 2i \end{bmatrix} \quad (3.79)$$

Substitution in eqn 3.37 yields immediately

$$\frac{d\Omega}{dt} = -\frac{3\mu J_2 R_0^2 \cos i \sin^2 \theta^*}{hr^3} \quad (3.80)$$

Since from eqn 1.18 $\dot{\theta} = h/r^2$, divide both sides of the equation by $\dot{\theta}$ and eliminate r by eqn 1.14

$$\frac{d\Omega}{d\theta} = -3J_2(R_0/p)^2 \cos i [1 + e \cos(\theta^* - \omega)] \sin^2 \theta^* \quad (3.81)$$

Take the average value of this equation over $(0, 2\pi)$ in θ^* resulting in

$$\left[\frac{d\Omega}{d\theta} \right]_{av} = -(3/2)J_2(R_0/p)^2 \cos i \quad (3.82)$$

or using N to denote *rates of change averaged over one revolution or a complete orbit*

$$\frac{d\Omega}{dN} = -3\pi J_2(R_0/p)^2 \cos i \quad (3.83)$$

The following results follow similarly

$$\frac{d\omega}{dN} = 3\pi J_2(R_0/p)^2 (2 - (5/2) \sin^2 i) \quad (3.84)$$

$$\frac{di}{dN} = \frac{dh}{dN} = \frac{de}{dN} = \frac{da}{dN} = 0 \quad (3.85)$$

Equation 3.83 indicates that an orbit will precess due to the oblateness of the Earth, ie. the line of the ascending node will retreat or advance depending on whether the inclination is less than or greater than 90 degrees. In the latter case it is known as a *retrograde* orbit. An important result arises if the inclination can be chosen so that the rate of precession is exactly equal to the angular rate of the Earth around the Sun, viz. 6.845×10^{-4} degrees per minute. In that case the orbit will remain invariant with respect to the Sun; it is known as a *Sun synchronous orbit*. This occurs when

$$(1/T) \frac{d\Omega}{dN} = 6.845 \times 10^{-4} \quad (3.86)$$

if $d\Omega/dN$ is in degrees per revolution and the orbital period T is in minutes. The latter is (for example) 90.52 minutes for a circular orbit at altitude 300 km. Using eqn 3.83 the above calculation yields the result that, at this altitude, the inclination should be 96.67 degrees.

3.4 Perturbations from the Sun and the Moon

Referring to Figure 3.1, the centre of mass of the Earth-Moon system is distance r_1 along the line

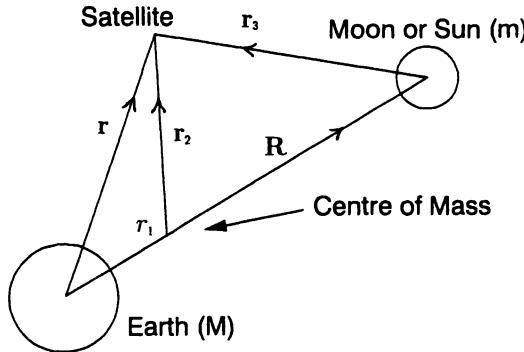


Fig. 3.1. Solar and Lunar Perturbations

from the Earth to the Moon (Sun). If M and m are the masses of the Earth and Moon (Sun) respectively

$$r_1/R = \nu/(1 + \nu) \quad (3.87)$$

where

$$\nu = m/M \quad (3.88)$$

The motion of the satellite relative to the centre of mass is given by

$$\ddot{\mathbf{r}}_2 = -\frac{\mu \mathbf{r}}{r^3} - \frac{\nu \mu \mathbf{r}_3}{r_3^3} \quad (3.89)$$

But

$$\mathbf{r}_3 = \mathbf{r} - \mathbf{R} \quad (3.90)$$

and

$$\mathbf{r}_2 = \mathbf{r} - \left(\frac{\nu}{1 + \nu} \right) \mathbf{R} \quad (3.91)$$

Therefore

$$\ddot{\mathbf{r}} - \left(\frac{\nu}{1 + \nu} \right) \ddot{\mathbf{R}} = -\frac{\mu \mathbf{r}}{r^3} - \frac{\nu \mu (\mathbf{r} - \mathbf{R})}{\|\mathbf{r} - \mathbf{R}\|^3} \quad (3.92)$$

However, given that the satellite mass is negligible

$$\ddot{\mathbf{R}} = -\frac{\mu(1 + \nu)\mathbf{R}}{R^3} \quad (3.93)$$

leading to the equation of motion of the satellite perturbed by the Moon (Sun).

$$\ddot{\mathbf{r}} = -\frac{\mu \mathbf{r}}{r^3} + \frac{\nu \mu (\mathbf{R} - \mathbf{r})}{\|\mathbf{R} - \mathbf{r}\|^3} - \frac{\nu \mu \mathbf{R}}{R^3} \quad (3.94)$$

The acceleration due to perturbations from the Moon (Sun) arises from the second and third terms. This equation can be employed for the Sun or the Moon but, because in the former case $R \gg r$, it is probably preferable to use for the Sun

$$\|\mathbf{R} - \mathbf{r}\|^3 \simeq R^3 \left(1 - \frac{3}{R^2} \mathbf{r} \cdot \mathbf{R} \right) \quad (3.95)$$

when

$$\ddot{\mathbf{r}} \simeq -\frac{\mu \mathbf{r}}{r^3} + \frac{\nu \mu}{R^3} \left[\frac{3}{R^2} (\mathbf{r} \cdot \mathbf{R}) \mathbf{R} - \mathbf{r} \right] \quad (3.96)$$

3.5 Effect of Atmospheric Drag

From eqn 2.27 the air velocity (excluding winds) is

$$\mathbf{v}_a = \mathbf{v} - \boldsymbol{\Omega}_e \times \mathbf{r} \quad (3.97)$$

As a standard result of aerodynamics the acceleration due to aerodynamic drag is

$$\mathbf{f} = -\frac{\rho S C_D}{2m} v_a \mathbf{v}_a \quad (3.98)$$

where ρ is the air density, m is the spacecraft mass, S the effective cross-sectional area, and C_D is the drag coefficient. It is difficult to calculate SC_D but the ballistic coefficient

$$B_c = \frac{m}{S C_D} \quad (3.99)$$

can be estimated from in-flight observations. It is typically in the range 25 to 100 kg per sq.m. Given a value for B_c

$$\mathbf{f} = -\frac{\rho v_a}{2B_c} \mathbf{v}_a \quad (3.100)$$

The air density must come from a model of the atmosphere which is also subject to uncertainties and variations, especially at the higher altitudes.

3.6 Illustrative Computations of Perturbations

The equations of section 3.2.6 have been programmed to permit simulation of perturbations from an oblate Earth, the Moon, the Sun, and atmospheric drag. The ballistic coefficient was taken to be 40 kg per sq.m and the model of the atmosphere was the same as that employed in chapter 2 (1962 U.S. Standard Atmosphere, Table 2.1, Ref. [45]) The orbits of the Moon and Sun have both been approximated as circular. The lunar orbit has an inclination to the ecliptic plane of 5.145 degrees with a sidereal period of 27.322 days. The longitude of the ascending node is approximately 178.7 degrees (estimated from 1997 values of the Right Ascension and Declination of the Moon, Ref. [39]). The simple coordinates of the lunar circular orbit have to be transformed first to ecliptic axes by Γ_2 of eqn 1.30, and then to Equatorial axes by

$$\begin{bmatrix} 1 & 0 \\ 0 & \cos \beta & -\sin \beta \\ 0 & \sin \beta & \cos \beta \end{bmatrix} \quad (3.101)$$

where β (the angle between the ecliptic and equatorial planes) is 23.45 degrees.

Example 3.6.1. Low Earth Orbit:

$$\begin{aligned} a &= 6678 \text{ km} & e &= 0.01497 & i &= 30.0 \text{ deg} \\ \omega &= 90.0 \text{ deg} & \Omega &= 0.0 \text{ deg} \end{aligned}$$

At this altitude (300 km) the perturbations from the Sun and Moon are negligible compared to those from the Oblate Earth. In agreement with the analytical results of section 3.3.3, the latter gives rise to

1. Zero average rate of change of all the orbital elements except ω and Ω .
There are very small periodic variations in a , e , and i .
2. Ω drifts -0.464 degrees per orbit.
3. ω drifts +0.727 degrees per orbit.

The effect of aerodynamic drag is significant, the height of the perigee being the critical parameter (200 km in this example). The computed reduction in the semi-major axis was 0.63 km per revolution. By continuing the calculation over a large number of orbits it was possible to estimate the lifetime of such a satellite (in the absence of corrections) as approximately 100 orbits. The latter is proportional to the ballistic coefficient and this result is in good agreement with data published elsewhere on the lifetime of satellite orbits (Figure 3-17, Ref. [57]).

3.7 Correcting Low Earth Orbits

In general the three components of a corrective velocity manoeuvre affect the 6 orbital elements, and therefore the six elements can be adjusted by making

two manoeuvres at different times in an orbit. It is not common to require the adjustment of all the orbital elements and consequently we take (as an example) the case of correcting a , e , and Ω . From the equations of section 3.2.5, if the manoeuvre is assumed to require expenditure of propellant over only a short interval of time,

$$\begin{aligned}\Delta e &= \frac{2}{v}(e + \cos \theta)\Delta V_t - \frac{r}{av} \sin \theta \Delta V_n \\ \Delta a &= \frac{2a^2 v}{\mu} \Delta V_t \\ \Delta \Omega &= \frac{r \sin \theta^*}{h \sin i} \Delta V_z\end{aligned}\quad (3.102)$$

where ΔV_t , ΔV_n , ΔV_z are the three components of the velocity correction.

Table 3.1. Correcting a Low Earth Near-Circular Orbit

Orbital error	Best angle	Velocity increment m/s
$\Delta \Omega = 1$ deg	$\theta^* = 90$ or 270 deg	$V_z = \pm 64.8$
$\Delta a = 10$ km	$\theta = 90$ or 270 deg	$V_t = 5.92$
$\Delta e = 0.001$	$\theta = 90$ or 270 deg	$V_n = \pm 7.78$

In order to see clearly what kind of correction is required for each orbital error, rearrange as

$$\begin{aligned}\Delta V_t &= \frac{\mu \Delta a}{2a^2 v} \\ \Delta V_n &= \frac{(e + \cos \theta) \mu \Delta a - a^2 v^2 \Delta e}{avr \sin \theta} \\ \Delta V_z &= \frac{h \sin i \Delta \Omega}{r \sin \theta^*}\end{aligned}\quad (3.103)$$

These equations have been employed to prepare Table 3.1 for a nominally circular orbit at 200 km altitude, inclination 28.5 degrees. The entries illustrate the velocity increments required to correct (one at a time) small errors in Ω , a , and e . Note how expensive it would be to offset the precession of an orbit due to the oblateness of the Earth.

3.8 Geostationary Orbits

3.8.1 Introduction

Geostationary or Geosynchronous orbits are to be assumed to be equatorial orbits with a period equal to the sidereal day (86164.1 seconds), ie. corresponding to the daily rotation of the Earth relative to the stars and not the

Sun. When the orbit is circular, at a radius of approximately 42164 km, the satellite will appear stationary to an observer on the Earth. The majority of communication satellites and some weather satellites are in such orbits and therefore, apart from their unique properties, they deserve special attention. The principal characteristics are discussed here but for more details readers should refer to a specialised text such as Pocha (Ref. [44]) or a comprehensive review paper with many references by Shrivastava (Ref. [49]).

Although the perturbations on satellites in geostationary orbits are very small, they become important due to the tight tolerances arising from mission requirements. For example, due to the narrow antenna beam widths of the ground transmitter, it is common with modern communication satellites to require that the satellite remains stationary relative to the ground within plus or minus 0.1 degree in both latitude and longitude. The term *station-keeping* is applied to the procedures for maintaining the satellite within the specified tolerances.

The simulation software already mentioned (and supplied via the publisher) is also applicable to geostationary orbits and has been applied to generate illustrative numbers below. However the orbits are singular both due to zero eccentricity and zero inclination. Zero eccentricity was handled by means of the set listed in section 3.2.6 and zero inclination can be circumvented by employing axes rotated through 90 degrees as explained in section 3.2.7. The only inconvenient result of this transformation is that the orbital elements i , and Ω have to be interpreted differently after the transformation. In the case of the equatorial geostationary orbit, if Δ denotes a perturbation, a small change of inclination of the orbit is actually calculated as the square root of $(\Delta i^2 + \Delta \Omega^2)$ from the transformed variables.

3.8.2 North-South Perturbations due to the Moon and Sun

If the inclination of the orbit drifts away from the Equator then, whatever the longitude of the observer, the satellite will appear to have a daily oscillation in latitude equal to the magnitude of the non-zero inclination. The changes in the inclination of a geostationary orbit arise from the effects of the Moon and the Sun.

Using the equations of section 3.2.6 with 90 degree rotated axes, the drift in inclination due to Lunar perturbations is shown in Appendix A.1 to be

$$\left[\frac{dinc}{dt} \right]_{av} = -\frac{3\mu_m \sin 2\epsilon_m}{8\omega_e R_m^3} \quad (3.104)$$

where the subscript m refers to the Moon and ω_e is the sidereal rate of rotation of the Earth. The rate of change has been averaged over a lunar period (approximately 28 days). The angle ϵ_m is the angle (18.3 degrees) between the lunar orbital and equatorial planes. This is calculated from the lunar orbital elements ($i_m = 5.15\text{deg}$, $\Omega_m = 178.7\text{deg}$ with respect to the ecliptic plane) and the inclination β of the equatorial plane to the ecliptic by

$$\cos \epsilon_m = \cos i_m \cos \beta - \sin i_m \cos \Omega_m \sin \beta \quad (3.105)$$

The magnitude of expression 3.104 is evaluated as 0.0367 degrees per 28 days or 0.478 degrees per year, these values being confirmed by computer simulation.

The perturbations from the Sun can be similarly calculated but it is inappropriate to average over one year. Computer simulations show that the change in inclination per 28 days varies from 0.0095 to 0.0390 degrees depending on the time of the year. The mean rate of change of inclination (quoted on an annual basis for comparison with lunar perturbations) is 0.319 degrees. Therefore the two effects are comparable but their combined effect varies a little from year to year depending on the phase of the Moon relative to the year.

In order to correct a non-zero inclination a manoeuvre perpendicular to the orbital plane is required. From eqn 3.53 for a velocity increment Δv_z

$$\Delta i = r \cos \theta^* \Delta v_z / h = \cos \theta^* \Delta v_z / v \quad (3.106)$$

Assume that the correction is applied at the best time. Since the velocity in geostationary orbit is 3075 m/s, it follows that the annual budget to correct lunar perturbations (0.478 deg/year) is 25.7 m/s and, to correct solar perturbations (0.319 deg/year) 17.1 m/s is required. Therefore the annual budget is 42.8 m/s but it will vary somewhat from year to year; see Pocha (section 6.3, Ref. [44]). If the latitude is to be maintained within plus or minus 0.1 degree, the corrections will be necessary at intervals of approximately 46 days.

3.8.3 East-West Perturbations due to Triaxial Effects

The principal sources of error in longitude are an East-West acceleration due to Triaxial gravitational terms and a departure from circularity due to the effect of solar pressure on the structure of the satellite. Note however that any perturbation such as the oblateness term that gives rise to only a *constant rate of change of longitude* is not a problem because it is compensated by a very small permanent offset in the altitude of the orbit. The triaxial terms are those dependent on longitude ϕ in eqn 3.63. Only the tangential disturbing acceleration is of concern here and, for a circular orbit at the equator, this is obtained from eqns 3.63 and 3.66 as

$$f_t = \frac{6\mu}{a^2} (R_0/a)^2 (S_2^2 \cos 2\phi - C_2^2 \sin 2\phi) \quad (3.107)$$

The higher order terms have a minor effect and have therefore been ignored. Substitution of numerical values from section 3.3.2 yields

$$(S_2^2 \cos 2\phi - C_2^2 \sin 2\phi) = -1.808 \times 10^{-6} \sin(2\phi + \pi/6) \quad (3.108)$$

Since the angular velocity in circular orbit is

$$\omega_e = \sqrt{\mu/a^3} \quad (3.109)$$

the rate of change of longitude due to a change in semi-major axis is by differentiation

$$\frac{d\dot{\theta}^*}{da} = -\frac{3\omega_e}{2a} \quad (3.110)$$

The perturbation of the semi-major axis is given by eqn 3.52 and consequently

$$\frac{d\dot{\theta}^*}{dt} = -\frac{2a^2v}{\mu} \cdot \frac{3\omega_e f_t}{2a} = -3f_t/a \quad (3.111)$$

Substitute eqns 3.107 and 3.108 to give the acceleration of longitude as

$$\frac{d\dot{\theta}^*}{dt} = 18\omega_e^2(R_0/a)^2 \times 1.808 \times 10^{-6} \sin(2\phi + \pi/6) \quad (3.112)$$

With the numerical values of (R_0, a, ω_e) we obtain finally

$$\frac{d\dot{\theta}^*}{dt} = 3.960 \times 10^{-15} \sin(2\phi + \pi/6) \quad (3.113)$$

in units of radians/sec². The magnitude is 1.68×10^{-3} in units of degree/rev² where rev denotes one complete orbit. Computer simulation is in agreement and the addition of higher order terms makes only a small difference to this result (Ref. [44]).

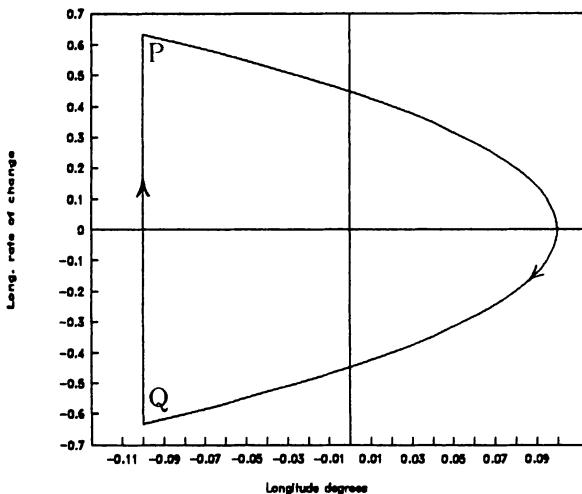


Fig. 3.2. East-West Stationkeeping

Consider now with the help of Figure 3.2 a strategy for maintaining the longitude within a given zone in the presence of a disturbing acceleration of magnitude c . Figure 3.2 is a diagram in which longitude rate $\dot{\phi}$ is plotted

against longitude ϕ ; $(\dot{\phi}, \ddot{\phi})$ are the xy coordinates. The following equations for time t have had constants adjusted so that the coordinates are $(-\phi_0, +\dot{\phi}_0)$ at zero time (point P) and $(-\phi_0, -\dot{\phi}_0)$ at time T_m (point Q).

$$\begin{aligned}\dot{\phi} &= cT_m/2 - ct \\ \phi &= -\phi_0 + cT_m t/2 - ct^2/2 \\ T_m &= 4\sqrt{c\phi_0/c}\end{aligned}\tag{3.114}$$

Assume that the angular rate is changed virtually instantaneously at point Q by the amount

$$+2\dot{\phi}_0 = +4\sqrt{c\phi_0}\tag{3.115}$$

The coordinates will have been restored to point P and thus a repetitive cycle develops in which the amplitude of the longitude is $\pm\phi_0$.

Let us now determine what manoeuvre is necessary to achieve this jump in the longitude rate. Equation 3.110 already relates change of longitude rate to change in semi-major axis, hence

$$\Delta a = -\frac{2a\dot{\phi}_0}{3\omega_e} = \frac{8a}{3\omega_e}\sqrt{c\phi_0}\tag{3.116}$$

Thus, it is required to change the semi-major axis while maintaining the eccentricity at zero. Assume that two equal tangential velocity increments Δv are applied, separated by an orbital angle of 180 degrees. Substitute in eqns 3.50, 3.51 and 3.52 for an approximately circular orbit.

$$\begin{aligned}\Delta e_x &= 2[\cos\theta^* + \cos(\theta^* + \pi)]\Delta v/v \\ \Delta e_y &= 2[\sin\theta^* + \sin(\theta^* + \pi)]\Delta v/v \\ \Delta a &= 4a^2v\Delta v/\mu\end{aligned}\tag{3.117}$$

It is evident that the change in eccentricity is zero for any angle θ^* and therefore the total velocity increment required is

$$2\Delta v = \frac{\mu\Delta a}{2a^2v} = (4/3)a\sqrt{c\phi_0}\tag{3.118}$$

having employed eqn 3.116 and because $v = \omega_e a$. From eqn 3.113, the worst case acceleration is $3.96 \times 10^{-15} \text{ rad/s}^2$ and if the tolerance ϕ_0 is taken as 0.1 degree, the sum of the two manoeuvres is calculated as 0.148 m/s. The interval between such corrections is given by eqn 3.114, viz. 30.7 days in the same worst case.

3.8.4 East-West Perturbations due to Solar Pressure

The effect of solar pressure on a satellite of mass m is to induce an acceleration away from the Sun. At the Earth's distance from the Sun the acceleration is approximately

$$f_s = 9.1 \times 10^{-6} A/m\tag{3.119}$$

in units of m/s^2 if the effective area A is in sq.m. and the mass in kg. The effective area can only be estimated although for a communication satellite equipped with large solar arrays, the area would arise largely from the latter. For example, with representative values of 80 sq.m. and a satellite mass of 1000 kg., the acceleration would be $7.28 \times 10^{-7} m/s^2$.

The effect of solar pressure is to render the geostationary orbit non-circular. The eccentricity vector (pointing at the perigee) rotates through a year so that it is always pointing towards the Sun. It is shown in Appendix A.2 that the increase in eccentricity per orbital revolution is approximately

$$\Delta e_y \simeq (3/2)f_s T/v \quad \Delta e_x = 0 \quad (3.120)$$

where T is the orbital period, f_s is the constant at eqn 3.119 and v is orbital velocity. By inserting numerical values for the above quoted representative case, we calculate (confirmed by computer simulation) that the eccentricity increases at the rate 3.06×10^{-5} per orbit.

To determine the effect as viewed by an observer, recall from section 3.2.6 that when f_z is zero and the eccentricity is small

$$\begin{aligned} \frac{d\theta^*}{dt} &= h/r^2 \simeq \sqrt{\mu/a^3}(1 + 2e_x \cos \theta^* + 2e_y \sin \theta^*) \\ &= \omega_e(1 + 2e_x \cos \theta^* + 2e_y \sin \theta^*) \end{aligned} \quad (3.121)$$

Substitute θ^* equal to $\omega_e t$ and let L denote longitude

$$\begin{aligned} \frac{dL}{dt} &= \frac{d\theta^*}{dt} - \omega_e = 2\omega_e(e_x \cos \omega_e t + e_y \sin \omega_e t) \\ &= 2\omega_e e \cos(\omega_e t - \omega) \end{aligned} \quad (3.122)$$

and integration yields

$$L = 2e \sin(\omega_e t - \omega) \quad (3.123)$$

i.e. the amplitude of the oscillation in longitude is equal to twice the eccentricity.

Let the rate of increase of eccentricity be e_r , then if the tolerance in longitude is ϕ_0 , it will be necessary to correct the eccentricity at intervals given by

$$T_m = \frac{\phi_0}{2e_r} \quad (3.124)$$

Using the value from the above example of the rate of increase of eccentricity (3.06×10^{-5} per orbit), a permissible error of 0.1 degree in longitude would require correction at intervals of 28.4 days. In order to correct this eccentricity two tangential manoeuvres $\pm \Delta v$ are employed separated by 180 degrees in orbit. Substitution in eqns 3.50 and 3.51 for a near circular orbit yields

$$\begin{aligned} \Delta e_x &= 4 \cos \theta^* \Delta v/v \\ \Delta e_y &= 4 \sin \theta^* \Delta v/v \end{aligned} \quad (3.125)$$

the change in major axis being zero. Therefore, to correct a non-zero value of e_y (eqn 3.120) the velocity increments should be applied at θ^* equal to 90 and 270 degrees. The sum of the two manoeuvres is

$$2\Delta v = ve_y/2 = v\phi_0/4 \quad (3.126)$$

eg. 1.34 m/s for a permitted longitude error of 0.1 degrees.

Thus, although the necessary increments of velocity to correct longitude errors are small, they are required at intervals of two or three months. In practice the strategy for correcting errors due to triaxiality and solar pressure would probably be integrated.

3.9 Rendezvous Guidance

3.9.1 Introduction

At an early stage of manned spaceflight the necessity arose to bring together in space two separate vehicles. This was an essential feature of the US mission Apollo to land a man on the Moon and consequently space rendezvous was first demonstrated in the preparatory programme between two Gemini vehicles in 1965. Since then there have been many examples of space rendezvous such as that between the American Apollo vehicle and the Soviet Soyuz 19 in 1975, and servicing of the Hubble Space Telescope by the US Space Shuttle Endeavour in 1992. In 1993 Russia and the United States scheduled future joint missions involving the Russian space station Mir and the Space Shuttle; rendezvous and docking will be a routine procedure in this programme.

The operation of rendezvous and docking is in reality two phases, (a) rendezvous guidance, and (b) proximity operations culminating in docking, ie. a controlled mating and securing of the two craft. The latter phase can be considered to start at a distance the order of 10 m and delicate attitude control of the two vehicles is required for the docking manoeuvre. This section is concerned only with the guidance phase from almost any other Earth orbit to within a separation in the range 10 to 100 m. Although space rendezvous has been practised since 1965 this section is based on later work at JPL published by Klumpp in 1987 (Ref. [30]) which is consistent with the availability of more powerful onboard computers and use of the Global Positioning System (GPS, appendix A.6). The latter permits rendezvous within one orbital period, since GPS can provide sufficiently accurate position and velocity of the *target* and *chaser* space vehicles at frequent intervals, in contrast to more ponderous orbit determination based on radar tracking from ground based sites.

In anticipation of regular manned rendezvous missions between the above mentioned Mir spacecraft and the Space Shuttle it became more important to obviate several preliminary orbits, to conserve propellant, and yet satisfy constraints with respect to safety and illumination. In devising a guidance

strategy, these criteria can be satisfied by (a) introducing the concept of a pre-computed *shaping trajectory* to satisfy the constraints in the approach phase, and (b) revised computation onboard of each of several manoeuvres by means of Lambert's algorithm. The resulting trajectories can be close to optimum in terms of propellant usage. The strategy is equally applicable to manned or unmanned flight although the safety criteria are likely to be less demanding in the latter case.

3.9.2 Guidance Strategy

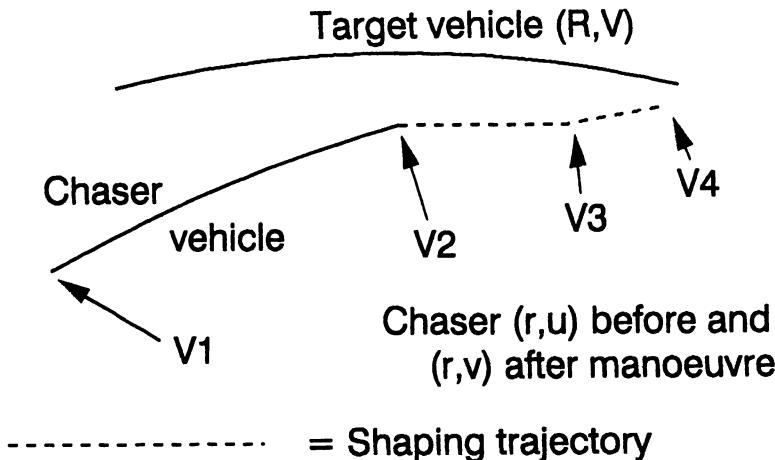


Fig. 3.3. Rendezvous Guidance

The strategy can be explained by reference to Figure 3.3 which illustrates the typical use of four impulse-type manoeuvres. The pre-computed shaping trajectory is shown as a broken line joining the three manoeuvre points (2,3,4). The last point is chosen close to the vehicle with a suitably small residual approach velocity, eg. 50 metres with a remaining velocity such that contact would occur after say 200 seconds. It is assumed that this final approach is under manual control or makes use of direct sensing of the target with reference to the chaser vehicle. The points of the shaping trajectory are chosen to satisfy the safety and illumination constraints. The necessary nominal manoeuvres are computed by working backwards from the last point applying Lambert's algorithm (section 1.5). It will be recalled that this algorithm permits calculation of the in-plane orbital elements given a starting position vector \mathbf{r}_1 , a finishing vector \mathbf{r}_2 , and the time of orbital transfer from point 1 to point 2. The out-of-plane elements Ω and i are obtained from the normal to the orbital plane ($\mathbf{r}_1 \times \mathbf{r}_2$); see section 1.3.2.

Referring again to Figure 3.3, the segment between points 1 and 2 is not pre-computed and in principle the starting point 1 can be any other Earth orbit. In practice the chaser orbit would be first modified (if necessary) so that both orbits are approximately coplanar. Let us now delineate the sequence of computations and manoeuvres that would be carried out in flight assuming that (1) the position and velocity of the chaser vehicle is available from onboard GPS, and (2) the orbit of the target vehicle is available corrected probably for perturbations from Earth oblateness. Navigation errors will of course arise and computed manoeuvres will be executed subject to errors in magnitude (shut-off) and pointing. Consequently, since *the logic of the following iterative sequence applies to all manoeuvres*, the shaping trajectory is not reproduced exactly. However, the effects of navigation and execution errors are limited as a result of the revision of the manoeuvres (in-flight) as updated determinations of position and velocity become available, ie. *this is a form of closed-loop guidance*.

1. Assume (at point 1) the latest orbit determination of the target vehicle, hence compute target position and velocity (\mathbf{R}, \mathbf{V}) at time T_f ahead, the provisional choice of time to allow for rendezvous equal to $(T_4 - T_1)$.
2. Step backwards to point 2 using Lambert's algorithm with the above revised (in-flight) values of \mathbf{R} and \mathbf{V} but with the pre-computed nominal manoeuvres V_4 and V_3 . The result is a revision of the position and velocity $(\mathbf{r}_2, \mathbf{v}_2)$ just after manoeuvre 2.
3. Given the determination of the current position and velocity $(\mathbf{r}_1, \mathbf{u}_1)$, apply Lambert's algorithm to the data \mathbf{r}_1 , \mathbf{r}_2 and time interval $(T_2 - T_1) = (T_f - T_4 + T_2)$. This yields the orbit between points 1 and 2 and consequently the required velocity \mathbf{v}_1 after manoeuvre 1.
4. Apply manoeuvre 1 equal to $(\mathbf{v}_1 - \mathbf{u}_1)$.

The above sequence is repeated for point 2 after the time $(T_2 - T_1)$ has elapsed, when updated values of position and velocity for both target and chaser are employed, and so forth for points 3 and 4. In the simplest form of the guidance strategy the time interval allowed for the shaping trajectory $(T_4 - T_2)$ is fixed but *there is scope for optimization, beforehand and in flight*, especially with respect to the time interval allowed for rendezvous T_f and the starting time T_1 . Klumpp (Ref. [30]) has elaborated the optimization that can be included in the software of an onboard computer.

3.9.3 Examples by Simulation

The above algorithm has been set up in software (available to the reader) to pre-compute the shaping trajectory and to simulate in- flight guidance including typical navigation and execution errors. The main program employs a number of modules some of which have already been applied with respect to earlier chapters, eg. orbital elements from position and velocity and vice versa,

solution of Kepler's equation, and Lambert's algorithm. However, in order to achieve the required accuracy bearing in mind the iterative structure of the last two routines, a double precision version of all routines was employed. In this connection, it should be mentioned that Klumpp (JPL) exploited a non-iterative solution of Lambert's algorithm due to Battin and Vaughan (Ref. [4]) although the author found solution by the secant iterative method perfectly adequate provided the special transformation of eqns 1.67 and 1.68 is included.

The details of one illustrative simulation are as follows, where conventional symbols for the orbital elements are in use.

Target Orbit : $a=6678$ km, $e=0.050$, $\omega=-90.0$ deg, $i=30.0$ deg, $\Omega=0.0$ deg, $\theta=0.0$ deg at zero time.

Initial Chaser Orbit : $a=6643$ km, $e=0.050$, $\omega=-90.0$ deg, $i=30.0$ deg, $\Omega=0.0$ deg, $\theta=-0.7354$ deg at zero time.

Time interval for rendezvous : $T_f=3300$ seconds.

One sigma errors in GPS Navigation : components of position and velocity (relative to target)= 20 m and 1 cm/s respectively.

One sigma errors in executing manoeuvres : components of velocity = 1 per cent of the magnitude of the velocity increment.

Table 3.2. Simulated Rendezvous without Errors

Time to go seconds	Position vector metres			Velocity vector m/s after manoeuvre		
	-80999	0	-33745	30.91	-8.15	99.50
500	200.0	700.0	0.0	-0.46	-1.36	-0.05
200	100.0	250.0	0.0	-0.38	-0.74	-0.00
0	32.0	95.0	0.0	-0.00	0.25	-0.45

Coordinates of position and velocity are quoted in Table 3.2 relative to the target vehicle in local horizontal axes at the specified time, ie. z is along the target local upwards vertical, x is East, and z is North (Figure 2.1). The transformation from these to inertial equatorial axes is given by eqn 2.77 if τ' is interpreted as Right Ascension and δ as Declination, calculable from the target position vector in equatorial axes at the appropriate time. Table 3.2 refers to a simulated rendezvous under ideal conditions, ie. no errors of navigation or manoeuvre execution. The approach starts from an orbit 35 km lower with the chaser behind at an initial separation of 87.7 km, and 3300 seconds is allowed for rendezvous at 100 metres from the target. *The shaping trajectory was chosen in order that the final approach is along the normal to the orbital plane.* This is merely an example of any final direction that can be selected depending on relevant constraints. By way of confirmation, the position vectors listed at lines 2,3,4 of Table 3.2 agree exactly with the

shaping trajectory specified as input to the program. Note that the spacecraft would ideally be left in a final state such that the remaining separation is closed within 200 seconds at a relative velocity of approximately 0.5 metres per second. The manoeuvre velocity increments (not listed in Table 3.2) were in order of application 13.0, 13.2, 0.79, and 1.16 m/s. The sum of 28.2 m/s compares with 20.2 m/s if an optimum Hohmann transfer were made between two coplanar *circular* orbits 35 km apart at this altitude; see eqns 1.94 and 1.97.

Table 3.3. Simulated Rendezvous with Random Errors

Time to go seconds	Position vector metres			Velocity vector m/s after manoeuvre		
	965.6	676.9	-232.8	-2.95	-1.17	0.71
500	-661.5	758.8	255.7	2.32	-1.43	-0.98
	-482.6	750.1	301.6	1.63	-1.48	-1.02
	-590.5	754.4	355.0	1.99	-1.40	-1.39
	102.5	283.5	-17.5	-0.38	-0.85	0.13
200	114.8	281.1	-11.8	-0.58	-0.94	-0.01
	66.7	266.5	14.0	-0.31	-0.83	-0.12
	93.2	289.9	-36.6	-0.42	-1.01	0.05
	33.7	106.3	8.6	-0.17	-0.51	0.11
0	8.2	87.6	-18.9	-0.05	-0.52	0.14
	16.2	94.7	-14.0	-0.04	-0.51	0.15
	21.5	80.4	-30.0	-0.11	-0.46	0.22

Table 3.3 illustrates the effects of typical navigation and execution errors, navigation position and execution errors being the more significant. Line 1 is omitted because it is the same as that in Table 3.2. The four lines for each entry "time to go" correspond to four different randomizations in the simulation of random errors of navigation and execution. The assumed errors clearly give rise to significant dispersion in the approach trajectory, although the approach from 200 seconds (time-to-go) is approximately along the orbit normal at the desired relative velocity. Note again that a further 200 seconds *after the above zero time-to-go* (at the closing residual velocity of 0.5 m/s) was assumed before proximity operations commenced.

3.9.4 Other references to Rendezvous and Docking

Other references of possible interest to the reader are as follows, although several academic papers more concerned with control theory are not included.

1. Pairet (Ref. [43] in French 1993): review of European developments, especially the practical aspects.
2. Serrano-Martinez (Ref. [48] with 23 references 1995): methodology of simulation tools for rendezvous and docking.

3. Griffin et alia (Ref. [23] 1988): radar for Shuttle rendezvous.
4. Saponaro (Ref. [47] 1970): Apollo guidance computer for rendezvous.

3.10 Software

The simulations of this section were carried out by means of C++ programs grouped together in the file RTH2.CPP which has been deposited in the Internet FTP server of the publisher as specified in the preface. The routines (all in double precision versions) included in that file are as follows.

Main : main program (1) to simulate perturbation of Earth orbits by integration of the equations in section 3.2.6 with axis change as necessary (section 3.2.7), and (2) to simulate rendezvous guidance based on Lambert's algorithm.

Transform : transformation between equatorial and other axes.

Sun : position vector of the Sun in equatorial axes.

Moon : position vector of the Moon in equatorial axes.

Runge_Kutta4 : see section 2.10.

RHS : right hand sides of state equations of section 3.2.6.

Earth : gravitational non-spherical field of the Earth.

Moon_perturb : perturbing acceleration due to the Moon.

Sun_perturb : perturbing acceleration due to the Sun.

elements_to_x : change from set of orbital elements to a state vector.

x_to_elements : change from a state vector to set of the orbital elements.

Orb_rv : transformation from orbital elements ($e_x, e_y, i, \Omega, \theta^*, r, h$) to vectors of position and velocity.

Trans_h_eq : given input position vector, calculate transformation from local horizontal to equatorial axes.

spacecraft_rv : given inputs of orbital elements of target and relative position and velocity, calculate position and velocity in equatorial axes.

backwards : propagate position and velocity backwards (using stored manoeuvres) from final position and velocity.

propagate : propagate position and velocity vectors from one time to another with no manoeuvres in the interval.

rv_rel_h : given position and velocity vectors at specified time, calculate position and velocity relative to target in local horizontal axes.

random_normal : approximation to a random normal distribution.

drag : (section 2.10) aerodynamic force as a function of the position and velocity vectors.

atmosphere : (section 2.10) US Standard Atmosphere 1962.

The files ORBIT2.CPP (section 1.9) and MATRIX2.CPP of orbit and matrix routines in double precision are also required.

CHAPTER 4

ORBIT DETERMINATION

4.1 Introduction

Before changing or correcting the trajectory of a spacecraft it is necessary to have determined the position and velocity of the vehicle relative to the Sun, the Earth or a planetary body, ie. navigation is a prerequisite for guidance. Since spacecraft are usually moving in a gravitational field due predominantly to one body such as the Earth or the Sun, the estimation of position and velocity is equivalent to the determination of the orbital parameters. Even when other gravitational sources are significant, we have seen that the use of osculating (varying) orbital elements is merely an alternative characterization to the use of the six components of position and velocity.

In determining orbits, three cases arise as follows.

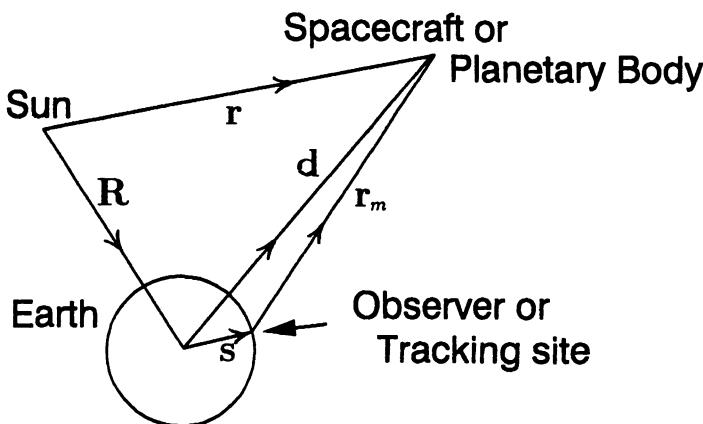


Fig. 4.1. Determining Planetary Orbits

1. The spacecraft is essentially in the gravitational field of the Sun (Figure 4.1) but the orbit is determined by radiometric measurements of range, range-rate and possibly angles from one or more tracking sites on the

Earth. It may also have been necessary to obtain or improve the orbit of a planetary target body such as a comet by astronomical observations from the Earth.

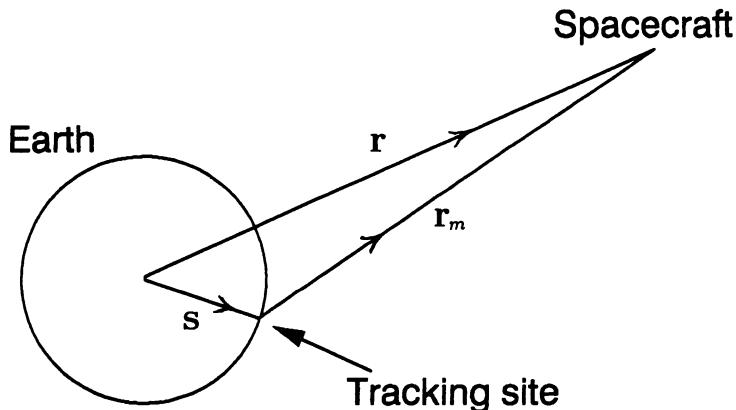


Fig. 4.2. Determining Earth Satellite Orbits

2. The spacecraft is an Earth satellite when the dominant gravitational body is the Earth. The orbit would again be determined by means of ground-based radiometric measurements or possibly by the Global Positioning System (see below).
3. The spacecraft is approaching another planetary body, viz. a planet, asteroid, or comet. Figure 4.1 applies to this case but eventually the gravitational field of the target body will dominate and it will be the orbit *relative to the target body* that becomes most relevant. Optical sightings (by means of an onboard camera) of the target against a star background will then provide crucial additional measurements.

The subject of orbit determination dates from ancient times, for example Ptolemy constructed an Earth centred model to fit observations of the inner solar system. In the sixteenth century Copernicus had the Sun at the centre with the planets orbiting in circles but the major advance was due to Kepler in the early seventeenth century. He exploited Brahe's observations to deduce that the orbits were ellipses with the Sun at one focus and discovered empirically his three laws of planetary motion; see the introduction to Chapter 1. Later in that century Newton showed that Kepler's empirical laws could be deduced from the *Law of Universal Gravitation*.

After Kepler, astronomers had all the *proportions* of the Solar System but not the *scale*. The first step was to establish the radius of the Earth which was estimated as early as 200 BC to within 6 per cent (but ignored by

Ptolemy) by means of a *shadow stick*. The shadow from a vertical stick was measured and a second point was established a known distance to the South, where the Sun was directly overhead. Given later refined determinations of the Earth's radius, the scale of the Solar System was first obtained by French observers in 1671 who measured the parallax of Mars, the apparent change in position relative to background stars due to the rotation of the Earth. The distance of the Moon was similarly an important quantitative parameter enabling Newton to derive Kepler's laws from his Law of Gravitation. In the eighteenth century the perturbation of planetary orbits was studied and the whole subject of orbit determination (based on astronomical observations) was greatly advanced by famous mathematicians such as Euler, D'Alembert, Lagrange, and Laplace. Although some of the material below derives from classical methods, the subject is developed to include (a) ground-based radar measurements of range and range-rate (doppler), and (b) modern computer-oriented methods for the refinement of orbits. The orbits of satellites in low Earth orbits will be determined more and more in the future by means of the Global Positioning System, therefore some explanation of GPS is given in appendix A.6.

4.2 First Estimates of Orbits

4.2.1 Transformation to Observables

Although orbits of the planets are well established we must specify methods for obtaining first estimates of orbits from a few preliminary observations, eg. a spacecraft shortly after separation from the launch system. For both approximate and refined methods of orbit determination it is necessary to have a transformation from the position and velocity vectors \mathbf{r} and \mathbf{v} (Figures 4.1 and 4.2) to the observables from a ground site, viz. angles of azimuth and elevation, range and range-rate. A computer routine for such a transformation would proceed as follows.

1. If \mathbf{r} and \mathbf{v} are in Earth-centred equatorial axes proceed to step 2, otherwise transform from Sun-centred ecliptic axes to Earth-centred equatorial axes, the rotational transformation being

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \beta & -\sin \beta \\ 0 & \sin \beta & \cos \beta \end{bmatrix} \quad (4.1)$$

where β is the inclination of the polar axis to the ecliptic plane. The position and velocity vectors of the Earth relative to the Sun (\mathbf{R} and \mathbf{V} as a function of time) will also be required. They can be generated from the orbital elements but since the orbit is nearly circular it is convenient to employ an available series formula as delineated in Appendix A.4.

2. Specify the position and velocity vectors of the observer in Earth-centred equatorial axes. Thus if the latitude and longitude of the observer are (θ_0, ϕ_0) put

$$\alpha = \phi_0 + \omega_e t + \phi'_0 \quad (4.2)$$

ω_e is the sidereal rate of rotation of the Earth (relative to the stars) and ϕ'_0 is a phase angle that must be chosen to give the correct position of the Sun at noon; it will also depend on the reference value of the time t . The position and velocity of the observer is given by

$$\mathbf{s} = r_0 \begin{bmatrix} \cos \theta_0 \cos \alpha \\ \cos \theta_0 \sin \alpha \\ \sin \theta_0 \end{bmatrix} \quad (4.3)$$

$$\dot{\mathbf{s}} = \omega_e r_0 \begin{bmatrix} -\cos \theta_0 \sin \alpha \\ \cos \theta_0 \cos \alpha \\ 0 \end{bmatrix} \quad (4.4)$$

3. Calculate the position and velocity of the object or spacecraft relative to the observer, viz. r_m and v_m of Figures 4.1 and 4.2. The observed range is r_m and the observed range-rate is the scalar product of v_m with the unit vector along r_m .
4. Test if the object can be seen by the observer, ie. it is more than the small angle ϵ above the horizon if

$$\mathbf{s} \cdot \mathbf{r}_m > r_0 r_m \epsilon \quad (4.5)$$

5. Transform \mathbf{r}_m to rotating equatorial axes where the first axis is through the observer site. This is a rotation matrix using the angle α from eqn 4.2, viz.

$$\tilde{\mathbf{r}}_m = \begin{bmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{r}_m \quad (4.6)$$

6. Transform to *observer axes* (the h axes of Figure 2.1) by means of

$$\mathbf{r}'_m = \begin{bmatrix} \sin \theta_0 & 0 & -\cos \theta_0 \\ 0 & 1 & 0 \\ \cos \theta_0 & 0 & \sin \theta_0 \end{bmatrix} \tilde{\mathbf{r}}_m \quad (4.7)$$

θ_0 being the observer's longitude.

7. The observed angle of elevation above the horizon is now calculated as

$$\theta' = \arcsin[r_m(3)/r_m] \quad (4.8)$$

and the azimuth angle East of North is

$$\phi' = \pi - \arctan[r_m(2)/r_m(1)] \quad (4.9)$$

4.2.2 Planetary Orbits

Consider the problem of estimating the planetary orbit of a distant object from only three ground-based astronomical angular sightings. First we must transform this data into three unit vectors (at the three different times) from the centre of the Earth to the object. This is the inverse of the operations of the previous section, viz.

1. Assuming a very approximate value for the distance from the observer to the object, transform the vector \mathbf{r}_m from observer to rotating equatorial axes,
2. Convert to Earth-centred Equatorial axes.
3. From the known observer position vector \mathbf{s} as a function of time, calculate the vector \mathbf{d} of Figure 4.1 by means of

$$\mathbf{d} = \mathbf{s} + \mathbf{r}_m \quad (4.10)$$

and the required direction is the unit vector corresponding to \mathbf{d} .

The planetary orbit can now be estimated from the three observed unit vectors ($\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$) at times (t_1, t_2, t_3). The method was proposed by Laplace (Ref. [3]). The important point to note is that *the orbital position and velocity vectors are obtained indirectly by exploiting the known equations of motion in the gravitational field of the Sun*, ie. assuming knowledge of the value of the solar gravitational constant and the Earth's orbit about the Sun.

Referring to Figure 4.1 and approximating the mass of the Sun as dominant, the equations of motion of the object and the Earth satisfy

$$\ddot{\mathbf{r}} + \frac{\mu}{r^3} \mathbf{r} = 0 \quad (4.11)$$

$$\ddot{\mathbf{R}} + \frac{\mu}{R^3} \mathbf{R} = 0 \quad (4.12)$$

where

$$\mathbf{r} = \mathbf{R} + \mathbf{d} \quad (4.13)$$

Put

$$\mathbf{d} = d\mathbf{i} \quad (4.14)$$

leading to

$$\begin{aligned} \ddot{\mathbf{d}} &= \ddot{\mathbf{r}} - \ddot{\mathbf{R}} \\ &= -\frac{\mu}{r^3}(\mathbf{R} + \mathbf{d}) + \frac{\mu}{R^3}\mathbf{R} \\ &= \frac{d}{dt} \left[d\dot{\mathbf{i}} + d \frac{d\mathbf{i}}{dt} \right] \end{aligned} \quad (4.15)$$

It follows that

$$\left(\ddot{d} + \frac{\mu}{r^3}d \right)\mathbf{i} + 2d\frac{d\mathbf{i}}{dt} + d\frac{d^2\mathbf{i}}{dt^2} = \mu(1/R^3 - 1/r^3)\mathbf{R} \quad (4.16)$$

Form the scalar product first with the vector $\mathbf{i} \times d\mathbf{i}/dt$ and second with $\mathbf{i} \times d^2\mathbf{i}/dt^2$, yielding

$$Ad = \mu(1/R^3 - 1/r^3)B \quad (4.17)$$

$$2Ad \frac{dd}{dt} = \mu(1/R^3 - 1/r^3)C \quad (4.18)$$

where

$$A = \left(\mathbf{i} \times \frac{d\mathbf{i}}{dt} \right) \cdot \frac{d^2\mathbf{i}}{dt^2} \quad (4.19)$$

$$B = \left(\mathbf{i} \times \frac{d\mathbf{i}}{dt} \right) \cdot \mathbf{R} \quad (4.20)$$

$$C = \left(\mathbf{i} \times \mathbf{R} \cdot \frac{d^2\mathbf{i}}{dt^2} \right) \quad (4.21)$$

The additional equation from the geometry is

$$r^2 = d^2 + R^2 + 2d(\mathbf{i} \cdot \mathbf{R}) \quad (4.22)$$

To proceed with a numerical solution, note that the Earth's position vector \mathbf{R} is known. Let the unit vector \mathbf{i} correspond to the middle observation \mathbf{i}_2 and approximate it about that time by the series

$$\mathbf{i}(t) = \mathbf{i}_2 + (t - t_2) \frac{d\mathbf{i}_2}{dt} + \frac{1}{2}(t - t_2)^2 \frac{d^2\mathbf{i}_2}{dt^2} \quad (4.23)$$

The derivatives can be evaluated from the adjoining unit vectors as follows

$$\frac{d\mathbf{i}_2}{dt} = [\mathbf{i}_3(t_1 - t_2)^2 - \mathbf{i}_1(t_3 - t_2)^2 - \mathbf{i}_2(t_1 - 2t_2 + t_3)(t_1 - t_3)] / D \quad (4.24)$$

$$\frac{d^2\mathbf{i}_2}{dt^2} = 2[-\mathbf{i}_3(t_1 - t_2) + \mathbf{i}_1(t_3 - t_2) + \mathbf{i}_2(t_1 - t_3)] / D \quad (4.25)$$

where

$$D = (t_1 - t_2)(t_3 - t_2)(t_1 - t_3) \quad (4.26)$$

Substitute eqn 4.22 into 4.17 yielding a single nonlinear equation for d which must be solved iteratively, eg. by the secant method (eqn 1.62). Even when restricted to positive values there can however be two solutions depending on the starting estimate for d although it is usually easy to reject one of a very small magnitude.

Given the numerical solution for the distance from the observer to the body it is possible to repeat the above transformations from observed angles to unit vectors (from the centre of the Earth), bearing in mind a rough estimate was assumed for that distance. This is not however critical because it affects only the parallax correction to the optical sightings. The value for

d permits the solution for the position vector of the object from eqn 4.13 in the form

$$\mathbf{r} = d\mathbf{i} + \mathbf{R} \quad (4.27)$$

Equation 4.18 can now be solved for dd/dt which in turn yields the velocity vector by differentiation of eqn 4.27

$$\mathbf{v} = \dot{\mathbf{r}} = d\frac{d\mathbf{i}}{dt} + \mathbf{i}\frac{dd}{dt} + \mathbf{V} \quad (4.28)$$

\mathbf{V} being the known velocity vector of the Earth relative to the Sun in Equatorial axes. These solutions for the position and velocity vectors, after transformation to Ecliptic axes, can now be employed to calculate the orbital elements as in section 1.3.2.

4.2.3 Computer Examples of Planetary Orbits

Two examples are included of calculating first estimates of a planetary orbit from astronomical observations. They represent a significant computer exercise for the student following the previous section although suitable software is available to the reader; see section 4.5. All the necessary constants are included in that software and the observer is defined to be using GMT at the author's location, viz. Bristol, UK latitude=51.467 degrees, longitude=357.417 degrees.

Example 4.2.1. Three sightings of the planet Mars were made at zero hours on the following dates: 25 April 1997, 4 May 1997, 14 May 1997. The angular data in degrees is listed below where all times are relative to zero hours 20 Feb 1996 (Julian date 2450133.5).

days	azimuth	elevation
430.0	232.854	34.084
439.0	241.365	29.099
449.0	248.856	23.413

Using Laplace's algorithm of section 4.2.2 the orbital parameters were estimated as follows where the true values are in brackets. The latter were obtained from tables of "Planetary Coordinates" for the interval 1995-2002 (p.313, Ref. [39]).

$a=1.538$ (1.523) AU, $e=0.0949$ (0.0934), time at perihelion = - 13.265 (0.0) days, $\omega=283.3$ (286.4) deg, inc = 1.9 (1.9) deg, $\Omega=49.2$ (49.6) deg.

The estimates are very good except that the perihelion is not well determined since the orbit is only slightly eccentric. Since double precision computations were in use, the errors arise almost entirely from the discrete approximations for the first and second derivatives of the unit vector from Earth to Mars, eqns 4.24 and 4.25. If we simulate more realistic random angular errors (with standard deviation 5 arc seconds) then the estimates vary somewhat as illustrated in Table 4.1 for different randomizations.

Table 4.1. First Estimates⁽¹⁾ of the Orbit of Mars

a AU	e	t_0 days	ω	inc	Ω
1.560	0.0956	-45.714	273.1	1.9	49.1
1.539	0.0950	-15.365	282.5	1.9	49.2
1.530	0.0957	-1.929	286.9	1.9	49.2
1.578	0.0959	-67.361	267.5	1.9	48.7
1.513	0.0983	17.770	292.5	1.8	49.5

¹ True values are above table in brackets.

Example 4.2.2. Three sightings of the comet Hale-Bopp were made in February 1997. The angular data in degrees is listed below where all times are relative to zero hours 1 Feb 1997 (Julian date 2450480.5).

days	azimuth	elevation
-9.75	90.143	14.560
0.250	90.633	20.573
10.250	89.118	26.499

The true values below are as quoted by Yeomans (JPL) in a web site devoted to that comet at the time. Since the eccentricity is so close to unity the perihelion is specified instead of the semi-major axis which is of course very large and ill determined.

perihelion $R_{min}=0.914$ AU, $e=0.9951$, time at perihelion $t_0= 59.142$ days, $\omega=130.6$ deg, $inc=89.4$ deg, $\Omega=282.5$ deg.

Using Laplace's algorithm of section 4.2.2 the orbital parameters were estimated as shown in Table 4.2 where the observed angles have been simulated as subject to a random error of 5 arc seconds (standard deviation).

Table 4.2. First Estimates⁽¹⁾ of the Orbit of Comet Hale-Bopp

R_{min} AU	e	t_0 days	ω	inc	Ω
0.898	0.9503	59.040	132.0	90.0	282.9
0.904	0.9797	58.507	131.1	89.8	282.8
0.909	0.9918	58.373	130.5	89.7	282.6
0.911	0.9962	58.336	130.4	89.6	282.6
0.902	0.9661	58.783	131.5	89.9	282.8
0.916	1.0138	58.037	129.8	89.5	282.4

¹ True values are quoted above the table.

4.2.4 Spacecraft Orbits

Orbits of spacecraft are normally determined from ground tracking data, principally range and range-rate from the doppler signal. Angular data is

also useful for Earth satellites but less so for deep space missions because the angular changes are too small. However, as shown in Appendix A.3, the doppler signal is modulated at the Earth's sidereal spin rate and this gives an indirect measure of the Right Ascension and Declination.

Consider now the problem of obtaining a quick first estimate of an orbit from two position fixes, ie. range, azimuth and elevation angles. The transformation from these observables to the position vector \mathbf{r} relative to the Earth or Sun was given in section 4.2.1, although because of the above remarks, this method applies mainly to the estimation of Earth satellite orbits. Assume then that we have available two position vectors \mathbf{r}_1 and \mathbf{r}_2 at times t_1 and t_2 . As in section 4.2.2, we avoid calculating velocity by differences but instead exploit the known equations of motion in the dominant gravitational field.

Use a series expansion for the position vector in time t beginning at t_1 .

$$\mathbf{r} = \mathbf{a}_0 + t\mathbf{a}_1 + t^2\mathbf{a}_2 + t^3\mathbf{a}_3 \quad (4.29)$$

By differentiation, the velocity vector is

$$\mathbf{v} = \mathbf{a}_1 + 2t\mathbf{a}_2 + 3t^2\mathbf{a}_3 \quad (4.30)$$

From the known equation of motion

$$\ddot{\mathbf{r}} = -g\mathbf{r} \quad (4.31)$$

where

$$g = \mu/r^3 \quad (4.32)$$

Differentiate eqn 4.30 again

$$-g\mathbf{r} = 2\mathbf{a}_2 + 6t\mathbf{a}_3 \quad (4.33)$$

Set t equal to zero at \mathbf{r}_1 and t equal to $\tau = (t_2 - t_1)$ at \mathbf{r}_2 , hence the solutions

$$\mathbf{a}_0 = \mathbf{r}_1 \quad (4.34)$$

$$\mathbf{a}_2 = -\frac{1}{2}g_1\mathbf{r}_1 \quad (4.35)$$

It remains to solve (only) for $\mathbf{a}_2 = \mathbf{v}_1$ from

$$\begin{aligned} \mathbf{r}_2 &= \mathbf{a}_0 + \tau\mathbf{a}_1 + \tau^2\mathbf{a}_2 + \tau^3\mathbf{a}_3 \\ -g_2\mathbf{r}_2 &= 2\mathbf{a}_2 + 6\tau\mathbf{a}_3 \end{aligned} \quad (4.36)$$

Eliminate \mathbf{a}_3 from eqn 4.36 and substitute for \mathbf{a}_0 and \mathbf{a}_2 to yield

$$\mathbf{a}_1 = \mathbf{v}_1 = [-1/\tau + g_1\tau/3]\mathbf{r}_1 + [1/\tau + g_2\tau/6]\mathbf{r}_2 \quad (4.37)$$

Given these solutions for \mathbf{r}_1 and \mathbf{v}_1 the orbital elements can be calculated as specified in section 1.3.2. This calculation would be especially useful in first estimating an orbit about the Earth from early sparse data from one ground site. First estimates of spacecraft orbits relative to the Sun would normally be obtained by adding the Earth's velocity about the Sun to the hyperbolic escape velocity vector.

4.2.5 Computer Example of Earth Satellite

This is another computer exercise requiring modules based on section 4.2.1 but included in the software made available (section 4.5).

Example 4.2.3. A spacecraft passes over a ground site at latitude=51.467 degrees, longitude=357.417 degrees. The following measured data becomes available (in the absence of errors) relative to zero hours 4 May 1997.

hours	azimuth	elevation	range km
1.450	201.02	41.26	677.624
1.500	132.05	13.33	1443.765

The true orbit is as follows:

$a=6878.0$ km, $e=0.0100$, time at perigee $t_0=-0.389$ hours, $\omega=45.0$ deg, $inc=50.0$ deg, $\Omega=125.0$ deg.

If random errors are simulated with standard deviations of 0.05 degrees and 30.0 m for angles and range respectively, then the estimates of the orbit are obtained for different randomizations as shown in Table 4.3.

Table 4.3. First Estimates⁽¹⁾ of an Earth Satellite Orbit

a km	e	t_0 hours	ω	inc	Ω
6878.4	0.0098	-0.385	46.1	50.0	125.0
6875.4	0.0103	-0.401	41.9	50.0	125.0
6872.7	0.0088	-0.392	44.1	50.0	125.0
6880.7	0.0098	-0.377	48.1	50.0	125.1
6876.0	0.0100	-0.397	42.9	50.0	125.0
6872.1	0.0094	-0.404	41.2	50.0	125.0

¹ True values are quoted above the table.

4.3 Refinement of Orbits

4.3.1 Batch Least-Squares Fitting

Spacecraft are normally subject to ground-based tracking at regular intervals from two or three sites around the globe; the received data will comprise telemetry to monitor the status of the vehicle, commercial and scientific observed data, and radiometric data to permit orbital updating. In the early stages of a mission near the Earth, all intervals of "visibility" may be exploited from all available tracking sites to refine the orbit determination, but infrequent updating may be acceptable later in the mission, eg. a communication satellite in geostationary orbit or a space probe during a long uneventful leg of a mission at great distance from the Earth. Nonetheless,

both in the early and mature stages of a mission many measurements are available and commonly employed to refine or update the determination of the orbital parameters to a high order of accuracy.

The conventional method of determining the values of a set of constant parameters which best fit the observed data is by the *method of least squares*. The principal parameters here are either the orbital elements or equivalently the vectors of position and velocity at some reference time, *the epoch*. Due to orbital perturbations and other possible non-deterministic disturbances to the trajectory, the parameters may not be strictly constant but we will approximate them as such for the initial analytical development. Let the vectors of parameters be \mathbf{p} but note that the number of components n may be more than six; for example we may wish to include estimation of the parameters of a perturbing body or thrust misalignments of the onboard rocket motor used for manoeuvres. It is assumed that a model has been developed to permit calculation of the observed variables for any set of numerical values of the parameters. In the simplest case this would entail the formulae already discussed for computing observed range, etc. as a function of the orbital elements.

Therefore we can write for the i 'th scalar measurement

$$y(i) = f(i, p) + e(i) \quad (4.38)$$

where $f()$ is a computable function and $e(i)$ is the error on the i 'th measurement. Note (1) the count i is over different kinds of measurements at more than one site and at different times, and (2) henceforth in these sections, vectors such as p are *not printed in bold type*. Least-squares fitting for m measurements becomes the minimization of the function

$$V(p) = \sum_{i=1}^m [y(i) - f(i, p)]^2 \quad (4.39)$$

by adjustment of the vector p , given a starting value \tilde{p} .

If a sufficiently good first estimate is available then local linearization of the nonlinear function $f()$ should be valid, the function $V(p)$ will be convex, and *linear least squares* fitting will yield a solution. Thus, take a first order expansion of eqn 4.38 about the vector estimate \tilde{p}

$$y(i) = f(i, \tilde{p}) + h(i)^T \delta p + e(i) \quad (4.40)$$

where

$$h(i)^T = \left[\frac{\partial y(i)}{\partial p_1}, \dots, \frac{\partial y(i)}{\partial p_n} \right] \quad (4.41)$$

Define the residuals of this provisional fit as

$$z(i) = y(i) - f(i, \tilde{p}) \quad (4.42)$$

then to a first order of approximation

$$z = H \delta p + e \quad (4.43)$$

where (for a total of m measurements) z and e are now m -component vectors of the residuals and errors, and H is an $m \times n$ matrix, the rows of which are the left-hand-side of eqn 4.41. The function 4.39 to be minimized can now be written in matrix notation as

$$V(\delta p) = (H\delta p - z)^T (H\delta p - z) \quad (4.44)$$

Differentiate with respect to δp_1 and set to zero. Differentiate with respect to δp_2 and set to zero and so forth to obtain n equations. These n equations are in matrix form

$$H^T H \delta p = Hz \quad (4.45)$$

or by inverting the $n \times n$ matrix, the solution for the change in the parameter vector is

$$\delta p = [H^T H]^{-1} H^T z \quad (4.46)$$

Eqn 4.46 is the solution to the linear least-squares fitting problem but in this case, because local linearization was employed, it may be necessary to repeat the above process after re-computing the matrix H of partial derivatives about the improved estimate of the vector p , ie. re-linearization.

When the process has converged, the z in eqns 4.42 and 4.46 is equivalent to the measurement errors in eqn 4.38. If z in eqn 4.46 is replaced by the vector of measurement errors e , the resulting δp can be regarded as the residual errors in the estimate arising from the measurement errors. Thus

$$\delta p \delta p^T = [H^T H]^{-1} H^T [ee^T] H [H^T H]^{-1} \quad (4.47)$$

Assume that the measurement errors are independent, have zero mean value and have been normalized to have unit standard deviations, ie. eqn 4.38 has been divided by the standard deviations of each kind of measurement error. It follows that the expected value of $[ee^T]$ is equal to the unit matrix. By taking expected values, eqn 4.47 then reduces to

$$E[\delta p \delta p^T] = G = [H^T H]^{-1} \quad (4.48)$$

G is the theoretical error-covariance matrix of the estimated vector p . The inverse of that matrix is $[H^T H]$ is known as the information matrix. The diagonal of the symmetric matrix G consists of the n theoretical variances of the errors in estimating the n components of the parameter vector.

4.3.2 Generation of Partial Derivatives

The above $m \times n$ matrix H is an array of the partial derivatives of the observables at a number of sample intervals with respect to the set of constant parameters being estimated. In the simplest case, these would be six orbital elements but if the spacecraft moves from one dominant gravitational field to another (as in the case of injection from an Earth to an interplanetary

orbit) then one might select the parameters to be estimated as the six initial components of position and velocity.

The position and velocity vectors have to be computed at all the sample intervals in order to generate the observables as a function of the parameter vector p . This will entail either propagation forward in time by means of the formulae of sections 1.3.1 and 1.3.2 or numerical integration of the equations of motion in say rectangular coordinates. Due to the complexity of the formulae and given the availability of modern computers working in double precision arithmetic, it is recommended that the partial derivatives are evaluated by carrying out the nominal and n separate forward propagations, in which small perturbations are applied to each of the n parameters (one at a time). Provided the computed observables are stored at each sample interval, the whole matrix H can be formed by the $(n + 1)$ propagations.

4.3.3 The Square Root Information Filter

As more and more data is employed to refine the determination of the parameter vector, the elements of G the covariance matrix of eqn 4.48) become smaller and smaller. More precisely the matrix $[H^T H]$ approaches singularity, *yet this matrix has to be inverted* and consequently the results can become unduly sensitive to errors such as the approximation of linearization, computer word length, etc. The situation is in fact quite common and can be made worse by trying to accommodate large numbers of measurements which might serve to determine well only a subset of all the parameters being refined.

An alternative formulation of least-squares fitting is therefore preferred which is much better conditioned numerically. It works with the square root information matrix R instead of the covariance matrix G . Methods are known for decomposing a real symmetric positive definite $n \times n$ matrix such as G as follows.

$$G = SS^T \quad (4.49)$$

The non-symmetric $n \times n$ matrix S (of rank n) is known as the *square root covariance matrix* and its inverse is the *square root information matrix*. Thus

$$S = R^{-1} \quad (4.50)$$

and by substitution in eqn 4.49

$$G = [R^T R]^{-1} \quad (4.51)$$

Notice that as a result of eqn 4.49, it is ensured that G is positive definite because for any non-zero vector x

$$\begin{aligned} x^T G x &= x^T S S^T x \\ &= z^T z = \|z\|^2 > 0 \end{aligned} \quad (4.52)$$

where $z = S^T x$.

The square root information filter (SRIF) is developed as follows. The function 4.44 to be minimized (least squares fitting) can be re-written as

$$V(\delta p) = \|H\delta p - z\|^2 \quad (4.53)$$

where z is an m component vector. A norm of a vector is unaffected by an orthogonal transformation applied to that vector. Therefore, choose one such transformation T so that

$$TH = \begin{bmatrix} \tilde{R} \\ 0 \end{bmatrix}, \quad Tz = \begin{bmatrix} \tilde{z} \\ e \end{bmatrix} \quad (4.54)$$

where T is $m \times m$ and \tilde{R} is upper triangular $n \times n$. This is the Householder transformation, a well established algorithm of linear algebra; see for example Bierman (chapter 4, Ref. [6]) but it is mathematically equivalent to the "QR transformation" as treated in texts on linear algebra (Ref. [51]). It is a standard routine of computer libraries and is included in the software of this monograph (section 4.5). Since the norm of eqn 4.53 is unchanged by the transformation, it follows by substituting eqn 4.54 in 4.53 that

$$\begin{aligned} V(\delta p) &= \|TH\delta p - Tz\|^2 \\ &= \left\| \begin{bmatrix} \tilde{R} \\ 0 \end{bmatrix} \delta p - \begin{bmatrix} \tilde{z} \\ e \end{bmatrix} \right\|^2 \\ &= \|\tilde{R}\delta p - \tilde{z}\|^2 + \|e\|^2 \end{aligned} \quad (4.55)$$

which is minimized when

$$\tilde{R}\delta p = \tilde{z} \quad (4.56)$$

and the sum of the squares of the residuals is

$$V_{min} = \|e\|^2 \quad (4.57)$$

The solution of eqn 4.56 is easy because the $n \times n$ matrix \tilde{R} is upper triangular. Given that solution, suppose now that *another set of data becomes available* corresponding to a new vector of measurements but again represented by eqn 4.43. The least squares solution for the old and new measurements is achieved by minimizing

$$\begin{aligned} &\|\tilde{z} - \tilde{R}\delta p\|^2 + \|z - H\delta p\|^2 \\ &= \left\| \begin{bmatrix} \tilde{z} \\ z \end{bmatrix} - \begin{bmatrix} \tilde{R} \\ H \end{bmatrix} \delta p \right\|^2 \\ &= \left\| T \begin{bmatrix} \tilde{z} \\ z \end{bmatrix} - T \begin{bmatrix} \tilde{R} \\ H \end{bmatrix} \delta p \right\|^2 \end{aligned} \quad (4.58)$$

where T is another Householder orthogonal transformation. As a result

$$T \begin{bmatrix} \tilde{R} \\ H \end{bmatrix} = \begin{bmatrix} \hat{R} \\ 0 \end{bmatrix} \quad (4.59)$$

where \hat{R} is again upper triangular. Thus if

$$T \begin{bmatrix} \tilde{z} \\ z \end{bmatrix} = \begin{bmatrix} \hat{z} \\ e \end{bmatrix} \quad (4.60)$$

the new estimate of δp is given by

$$\hat{R}\delta\hat{p} = \hat{z} \quad (4.61)$$

i.e. requiring again the easy inversion of an upper triangular matrix. The sum of the squares of the residues is equal to

$$\|e\|^2 \quad (4.62)$$

Note that this is the sum of the squares of the residuals of the first and second set of data, a unique property of this algorithm.

The error-covariance matrix was derived at eqn 4.48. In the same way but using eqn 4.61 we calculate this matrix as

$$G = [\hat{R}^T \hat{R}]^{-1} \quad (4.63)$$

By comparison with the definition of the information matrix at eqn 4.51, it is evident that the matrix \hat{R} arising from the Householder transformation *is the square root information matrix*. The matrix \tilde{R} and vector \tilde{z} arose above from processing a previous set of data. We could alternatively regard them as *a priori* estimates, make \tilde{z} a vector of zero residuals, and calculate \tilde{R} from a matrix decomposition

$$\tilde{R}^T \tilde{R} = G_0^{-1} \quad (4.64)$$

where G_0 is the *a priori* covariance matrix. In practice it would usually be sufficient to treat G_0 as diagonal and set the initial \tilde{R} as diagonal with elements equal to the reciprocals of the standard deviations of the errors in the first estimates of the parameters. In the absence of any *a priori* information \tilde{R} and \tilde{z} are both set to have zero elements.

The calculation by means of SRIF of a correction δp to the parameter vector can be summarized as follows.

1. Assume the data set embodied in \tilde{R} and \tilde{z} either from *a priori* values or from a previous step of the SRIF algorithm with an earlier batch of data.
2. Include the new data by forming the vector of residuals z , and compute the matrix H of partial derivatives in eqn 4.43.
3. Apply the Householder algorithm to yield \hat{z} and \hat{R} using eqns 4.59 and 4.60.
4. Calculate the estimate δp by inversion of the upper triangular matrix in eqn 4.61.

In the case of a nonlinear system more than one iteration with the same data might be necessary.

All the available data would not normally be processed in one batch. Due to perturbations not included in the model, orbital parameters will change

slowly. Some bias terms may also drift in value. The size of the batch is consequently chosen large enough to permit good estimates but not so large that data is over a time interval that is too long. Use of a fresh set of data allows therefore updating of the parameter vector.

4.3.4 Bias Terms and the Consider Matrix

Equation 4.38 above relates the vector of constant parameters being estimated to the observed variables. The errors that might arise due to an imperfect model of the equations of motion, or in the measurements themselves, was approximated by the additive errors $e(i)$. The latter were assumed to be independent with zero mean value. The assumption of unbiased errors is usually unrealistic if very accurate estimates are being generated. In such cases the theoretical standard deviations of the errors in the parameter estimates (from the covariance matrix) decay to magnitudes which are too small; they become inconsistent with the actual errors of estimation. The corresponding square root information matrices R have become too large such that new data has little effect on the estimated values of the parameters.

The theoretical standard deviations of the estimates constitute a very useful complement to the actual estimates, both operationally and in simulation studies before a mission. It is desirable therefore to correct the covariance matrix when bias terms are expected in the errors $e(i)$. Introduce therefore additive bias terms into eqn 4.43 by means of a vector b of n_b components.

$$z = H\delta p + Bb + e \quad (4.65)$$

where B is an $m \times n_b$ matrix of partial derivatives generated by numerical perturbations in the same way as H . Now, the essential step of the above SRIF algorithm can be summarized by the following matrix operation.

$$T \begin{bmatrix} \tilde{R} & \tilde{z} \\ H & z \end{bmatrix} = \begin{bmatrix} \hat{R} & \hat{z} \\ 0 & e \end{bmatrix} \quad (4.66)$$

i.e. Householder's transformation T is applied to generate the upper triangular matrix \hat{R} and the vector \hat{z} . If the first order expansion of the measurement-model equation (modified for bias terms) is eqn 4.65, then the modified SRIF can be summarized as the operation

$$T \begin{bmatrix} \tilde{R}_p & \tilde{R}_{pb} & \tilde{z}_p \\ H & B & z \end{bmatrix} = \begin{bmatrix} \hat{R}_p & \hat{R}_{pb} & \hat{z}_p \\ 0 & \hat{B} & \hat{z} \end{bmatrix} \quad (4.67)$$

The first row of block matrices in 4.67 corresponds to

$$\hat{R}_p \delta p + \hat{R}_{pb} b = \hat{z}_p \quad (4.68)$$

where \hat{R}_p is upper triangular. Given that we are not attempting to solve for components of the bias vector, but our best estimate of it is the zero vector, the solution for the correction to the parameter vector is

$$\delta p = \hat{R}_p^{-1} \hat{z}_p \quad (4.69)$$

Even though the notation is slightly different to that in section 4.3.3, *the solution for δp as a function of the observed data is unchanged*. Hence we might refer to eqn 4.69 as the bias-free solution. However, we are now in a position to calculate the correction to the covariance matrix for bias terms. Assume that the necessary number of SRIF iterations with re-linearizations have occurred; then we can use eqn 4.68 to relate measurement and model errors e (instead of \hat{z}_p) and bias b to corresponding errors in estimating the parameters. Thus

$$\delta p \delta p^T = \hat{R}_p^{-1} (e - \hat{R}_{pb}) (e^T - b^T \hat{R}_{pb}^T) \hat{R}_p^{-T} \quad (4.70)$$

Take expected values of eqn 4.70 assuming that errors e and b are independent, whereupon the total error covariance matrix is given by

$$G' = [\hat{R}_p^T \hat{R}_p]^{-1} + \bar{S} G_b \bar{S}^T \quad (4.71)$$

where

$$\bar{S} = -\hat{R}_p^{-1} \hat{R}_{pb} \quad (4.72)$$

The first term of eqn 4.71 is the same as eqn 4.63 except for use of the subscript p ; it is *the bias-free covariance matrix*. The second term *represents the increase in the covariance terms due to the effect of bias errors and is called the consider matrix*. The matrix G_b is the (usually unchanging) symmetric covariance matrix of bias errors based on *a priori* information, eg. a diagonal matrix of estimated variances. Despite the usefulness of eqn 4.71, notice that the actual estimates would be unchanged and they have not been corrected for bias errors, but our calculation of the expected variances has been modified and they should not be too optimistic. Sometimes it may be feasible to include certain unknown constants in the estimation process, in which case the parameter vector is merely augmented by such constants.

4.4 Sequential Estimation

4.4.1 Non-constant Parameters

In some cases non-deterministic disturbances to the system are so significant that it is unrealistic to treat the problem as the estimation of a set of constant parameters. In this case it is more convenient to deal with the state equations of the system, ie. a set of first order differential equations. In the absence of stochastic (random) disturbances let the following equation describe the result of integrating the state equations from time t_k to t_{k+1} , where x is the vector of n states and b is the vector of n_b bias terms.

$$x(k+1) = \phi(x(k), b) \quad (4.73)$$

Assume now that any stochastic disturbances are sufficiently small that they can be approximated as additive, ie. eqn 4.73 becomes

$$x(k+1) = \phi(x(k), b) + C(k)w(k) \quad (4.74)$$

where w is a vector of n_w independent random inputs, each component of which has zero mean value and unit variance. Notice that this formulation permits the inclusion of parameters or disturbances which drift at random or are correlated (assuming that such terms are to be estimated). Thus a slowly varying parameter could be represented as the j 'th state equation in the form

$$x_j(k+1) = x_j(k) + \sigma\sqrt{h}w(k) \quad (4.75)$$

where h is the sub-interval of time between samples. The above corresponds to a *random walk* in which the variance would increase at the rate σ^2 per unit time.

It will be recalled that the treatment of nonlinear parameter estimation assumed that a first order expansion of the model was in use and the estimation algorithm referred to *corrections* to the previous best estimate of the parameter vector. In the same way assume a first order expansion of eqn 4.74 but, in order to retain a compact notation, we will continue to employ x as the state vector being estimated. Thus the state equations linearized about the best estimate are taken to be

$$x(k+1) = A_x(k)x(k) + A_b(k)b + C(k)w(k) \quad (4.76)$$

The matrices $A_x(k)$ and $A_b(k)$ have elements which are partial derivatives of elements of $x(k+1)$ with respect to elements of $x(k)$ and the bias vector b respectively, eg. the element (i, j) of $A_x(k)$ is

$$\frac{\partial x_i(k+1)}{\partial x_j(k)} \quad (4.77)$$

These matrices would be generated numerically in practice by n integrations over each subinterval of time when elements of $x(k)$ or b are perturbed one at a time.

4.4.2 Time Update with Sequential SRIF

The Square Root Information Filter has been developed above to update estimates of constant parameters as fresh measurements become available. SRIF can be applied to sequential estimation of non-constant parameters but

1. the batch will consist of only one or a small number of measurements at a given instant k , the *measurement update*.
2. after the measurement update, the information matrix \hat{R} and the estimated state vector \hat{x} must be propagated forward from time instant k to $k+1$, the separate *time update*.

The algorithm is quoted here without proof, having been adopted from Bierman (section 7.3, Ref. [6]). In that text, correlated process noise was denoted by the separate vector p but such terms are assumed here to be included in the state vector x . It will be recalled from eqn 4.67 that a measurement update gave rise to \hat{R}_p , \hat{R}_{pb} , and \hat{z}_p ; these are now denoted as $\hat{R}_x(k)$, $\hat{R}_{xb}(k)$, and $\hat{z}_x(k)$. Form the matrix of $(n_w + n)$ rows and $(n_w + n + n_b + 1)$ columns

$$S = \begin{bmatrix} I & 0 & 0 & 0 \\ -\bar{R}_x C(k) & \bar{R}_x & (\hat{R}_{xb}(k) - \bar{R}_x B_b) & \hat{z}_x(k) \end{bmatrix} \quad (4.78)$$

where

$$\bar{R}_x = \hat{R}_x(k) A_x(k)^{-1} \quad (4.79)$$

Apply a triangularization to the matrix S such that

1. the upper left $n_w \times n_w$ block is upper triangular
2. the lower left $n \times n_w$ block is zero.

The result is as follows although only the lower right $n \times (n + n_b + 1)$ block is of interest.

$$\hat{S} = \begin{bmatrix} R_w & R_{wx} & R_{wb} & \hat{z}_w(k+1) \\ 0 & \hat{R}_x(k+1) & \hat{R}_{xb}(k+1) & \hat{z}_x(k+1) \end{bmatrix} \quad (4.80)$$

The lower row of block matrices corresponds to the equation

$$\hat{R}_x(k+1)\hat{x}(k+1) + \hat{R}_{xb}(k+1)b = \hat{z}_x(k+1) \quad (4.81)$$

Given that the bias vector b is not being estimated and its expected value is zero, the solution for the estimated state vector after the time update is therefore

$$\hat{x}(k+1) = \hat{R}_x(k+1)^{-1}\hat{z}_x(k+1) \quad (4.82)$$

The required bias-free information matrix is $\hat{R}_x(k+1)$ and (as similarly calculated at eqn 4.71) the total covariance matrix after the time update is

$$P(k+1) = \left[\hat{R}_x(k+1)^T \hat{R}_x(k+1) \right]^{-1} + \bar{S} G_b \bar{S}^T \quad (4.83)$$

where

$$\bar{S} = -\hat{R}_x(k+1)^{-1}\hat{R}_{xb}(k+1) \quad (4.84)$$

The second term of eqn 4.83 is the consider matrix, ie. the increase due to bias terms. The above procedure, which again exploits the Householder algorithm, is the method of implementing the time update with SRIF when bias terms and stochastic inputs are present.

4.4.3 Measurement Update with sequential SRIF

The measurement update, when using SRIF for sequential estimation, is essentially as specified in section 4.3.4. It is merely necessary to allow for change of notation in the sequential case.

Thus instead of eqn 4.65 the measurement equation is

$$z(k) = H(k)x(k) + B(k)b + e(k) \quad (4.85)$$

where $z(k)$ is the vector of residuals (measurement less computed nominal measurement as at eqn 4.42) and $x(k)$ is the state vector relative to the current best estimate. *All the subsequent results of section 4.3.4 apply* by merely substituting $x(k)$ for δp , $\bar{R}_x(k)$ instead of \bar{R}_p , etc. By means of the triangularization of eqn 4.67 we obtain the measurement- updated bias-free information matrix $\hat{R}_x(k)$, the state vector $\hat{x}(k)$, and the consider matrix for bias errors in the measurements.

4.4.4 Sequential Estimation with Covariance Filters

Sequential estimation of a system subject to stochastic inputs can alternatively be carried out by propagating the covariance or the square root covariance matrix. The former is known as a *Kalman Filter* and is a well known piece of modern control theory. This algorithm is quoted without proof below because the subject is adequately treated in many texts such as Bryson and Ho (section 12.4, Ref. [8], although Kalman's original paper (Ref. [28]) is well worth attention.

The same notation is assumed as for the sequential SRIF estimator, ie. the linearized state equations are as at eqn 4.76 and the linearized equations for the measurement residuals are as at eqn 4.85. First, bias terms are excluded.

1. The time update to propagate the estimate $\hat{x}(k)$ and the covariance matrix $G(k)$ from instant k to instant $k + 1$ is

$$\tilde{x}(k+1) = A_x \hat{x}(k) \quad (4.86)$$

$$P(k+1) = A_x(k)G(k)A_x(k)^T + C(k)C(k)^T \quad (4.87)$$

given that the random inputs $w(k)$ to eqn 4.76 are independent, with zero mean value and unit variance. Equation 4.86 is based on a first order expansion, but in practice this would be better implemented by integration of the nonlinear state equations of the system.

2. The measurement update is

$$\hat{x}(k) = \tilde{x}(k) + K(k) [z(k) - H(k)\tilde{x}(k)] \quad (4.88)$$

$$K(k) = P(k)H(k)^T [H(k)P(k)H(k)^T + I]^{-1} \quad (4.89)$$

$$G(k) = P(k) - K(k)H(k)P(k) \quad (4.90)$$

Remember that $z(k)$ is the vector of residuals calculated from the actual measurements less the values computed with the nonlinear measurement function using the nominal state vector. The above $\tilde{x}(k)$ arises from a first order expansion about that nominal vector. $P(k)$ and $G(k)$ are the covariance matrices after time and measurement updates respectively.

Assume now that bias terms arise as in eqns 4.76 and 4.85. The bias terms are not estimated but the correction to the covariance matrix (the consider matrix) is required. Friedland (Ref. [21]) derived the following results which are equivalent to those of Bierman (Ref. [6]). Assume that the $(n \times n_b)$ matrix V_b is initiated as zero.

1. At the time update

$$U_b(k+1) = A_x(k)V_b(k) + A_b(k) \quad (4.91)$$

and the consider matrix at instant k is

$$U_b(k)G_bU_b(k)^T \quad (4.92)$$

2. At the measurement update

$$V_b(k) = U_b(k) - K(k)[H(k)U_b(k) + B(k)] \quad (4.93)$$

and the consider matrix at instant k is

$$V_b(k)G_bV_b(k)^T \quad (4.94)$$

In order to exploit the better numerical properties of the square root covariance matrix (eqn 4.49), a decomposed form of the Kalman Filter is available. See the UD covariance filter in Appendix A.5.

4.4.5 SRIF Example with an Earth Satellite Orbit

The refinement of orbit determination by means of SRIF in the batch mode (section 4.3.3) has been implemented in C++ software as summarized below in section 4.5. The allowance for bias matrices (section 4.3.4) has also been included.

The example is that of a spacecraft in a geostationary transfer orbit (GTO), the parameters being typical of an Ariane 4 launch from the near-equatorial launch site of the European Space Agency. The following ground tracking sites were simulated to be available, ie; latitude and longitude in degrees for each site.

Carnarvon	-24.90	113.72
Kourou	5.25	-52.80
Villafranca	40.45	-3.95
Malindi	-3.00	40.32
Redu	50.00	5.13

The true orbital elements were (in conventional notation)

a 24465.6 km
 e 0.7311
 θ_0 183.0 deg at zero time
 ω 178.0 deg
 i 7.0 deg
 Ω 0.0 deg

No perturbations such as those due to oblateness or atmospheric drag were simulated and consequently the above six parameters remain constant.

Fitting to observed data at 30 minute intervals over only approximately one orbit (10.6 hours) was employed because in reality atmospheric drag would reduce the semi-major axis by about 12 km every orbit. Thus SRIF would probably be employed with a batch of data corresponding to one orbit, and then the orbit updated with a fresh batch of data for the next orbit. As a result of the limitations of a typical onboard transponder, measurements from only one ground station at a time was assumed and the spacecraft was treated as "visible" only if it appeared at least 5 degrees above the horizon. Over the one orbit simulated here at 30 minute intervals, the spacecraft was visible 20 times from at least one ground site.

The following typical measurement errors (standard deviations) were assumed

measurement	random error	bias error	units
angles	0.02	0.02	degrees
range	20.0	20.0	m
range-rate	2.0	2.0	mm/s

and the initial conditions were as listed in Table 4.4. In this case the "sig" in the fourth column denotes the assumed *a priori* standard deviations of the first estimates of the six parameters.

Table 4.4. SRIF fit to Geostationary Transfer Orbit: First Estimates

Parameter	estimate	error	error/sig ⁽¹⁾
e_x	-0.73000	0.00068162	0.014
e_y	0.03000	0.00448404	0.090
θ^* deg	180.0	-3.00000	-0.150
p km	11400.00	12.38688	0.014
inc deg	9.00	2.00000	0.400
Ω deg	2.00	2.00000	0.400

¹ "sig" refers to the standard deviations of the first estimates.

Five or six re-linearizing iterations were necessary to achieve convergence with one batch of data. Table 4.5 shows the resulting estimates after six iterations when employing measurements of only range and range-rate (doppler) from any one station.

Table 4.5. Geostationary Transfer Orbit:
Final SRIF Estimates using range and range-rate

Parameter	estimate	error	error/sig ⁽¹⁾
e_x	-0.73069	-0.00001321	-0.894
e_y	0.02517	-0.00034177	-0.910
θ^* deg	183.03	0.02592	0.919
p km	11387.56	-0.05206	-0.728
inc deg	7.00	-0.00060	-0.871
Ω deg	-0.03	-0.02653	-0.917

¹ "sig" is the theoretical standard deviation corrected for bias terms.

Table 4.6. Geostationary Transfer Orbit:
Final SRIF Estimates using range, range-rate, and angles

Parameter	estimate	error	error/sig ⁽¹⁾
e_x	-0.73068	0.00000463	0.173
e_y	0.02562	0.00010216	0.153
θ^* deg	182.99	-0.00671	-0.133
p km	11387.66	0.04194	0.330
inc deg	7.00	-0.00134	-1.871
Ω deg	0.01	0.00770	0.149

¹ "sig" is the theoretical standard deviation corrected for bias terms.

The RMS residuals decay from 6593 to 0.84; the final value is very approximately unity, as it should be since the residual SRIFs and measurements have been scaled by dividing by the corresponding standard deviation of the measurements. Notice that the ratio *error/sig* is within the range ± 2 , confirming that the errors of this simulated estimation are consistent with the theoretical standard deviations corrected for bias terms (section 4.3.4). In analytical mission studies, parametric analysis is commonly carried out employing only the theoretical standard deviations. However, it is a wise precaution to check the convergence by randomized simulations (as above) since problems in achieving convergence can easily occur due to nonlinearities in the system, especially if the first estimates are not sufficiently accurate and *the set of parameters being estimated is barely observable*, ie. there are not enough sufficiently different measurements to determine the parameters. In the latter case one or more of the diagonal elements of the matrix \hat{R} (after the Householder transformation at eqn 4.59) tends to zero; see Lawson and Hanson (Ref. [32]).

Table 4.6 refers to a simulation which differs from that of Table 4.5 only because ground tracking angles (azimuth and elevation) have also been included in the measurement set. It is observed that the estimates are further

improved although in practice the determination based on only ranging would probably be acceptable.

4.4.6 SRIF Example for a Spacecraft Approaching a Comet

This example is that of a spacecraft approaching a comet when the latter becomes detectable by means of an onboard camera. See for example Hechler's early description of the *Rosetta* mission (Ref. [26]) and Noton (Ref. [41]). The comet is displayed against a background of stars which permits the line of sight from spacecraft to comet to be determined accurately as a known direction in space. This data is transmitted back to Earth and combined with the ground-based measurements of range and range-rate to allow an estimation of both the spacecraft orbit (relative to the Sun), and the position and velocity of the spacecraft *relative to the comet*. The twelve virtually constant parameters to be determined are therefore the heliocentric orbital elements of the spacecraft and the six components of position and velocity of the spacecraft relative to the comet at some reference time. This choice is equivalent to calculating the heliocentric orbital elements of the comet because the relative position vector is

$$\mathbf{r} = \mathbf{r}_{sc} - \mathbf{r}_c \quad (4.95)$$

where \mathbf{r}_{sc} and \mathbf{r}_c are the heliocentric position vectors of the spacecraft and comet respectively.

Consider first how the camera measurements are related to the parameters to be determined. Let the line of sight of the camera point in the direction (l, m, n) , ie. a unit vector with components in Ecliptic axes. Use this direction as the third axis of a set of *camera axes* and put

$$\mathbf{g}_3 = (l, m, n) \quad (4.96)$$

Take the first axis in the ecliptic plane and perpendicular to \mathbf{g}_3

$$\mathbf{g}_1 = (-m/\sigma, +l/\sigma, 0) \quad (4.97)$$

where

$$\sigma = \sqrt{l^2 + m^2} \quad (4.98)$$

Since

$$\mathbf{g}_2 = \mathbf{g}_3 \times \mathbf{g}_1 \quad (4.99)$$

$$\mathbf{g}_2 = (-nl/\sigma, -nm/\sigma, +\sigma) \quad (4.100)$$

It follows that the transformation from Ecliptic to camera axes is

$$\begin{aligned} S_1 &= [\mathbf{g}_1 \ \mathbf{g}_2 \ \mathbf{g}_3]^T \\ &= \begin{bmatrix} -m/\sigma & +l/\sigma & 0 \\ -nl/\sigma & -nm/\sigma & 0 \\ l & m & n \end{bmatrix} \end{aligned} \quad (4.101)$$

The direction cosines (l, m, n) would be calculated from the camera angular coordinates of known stars in the field of view of the camera. The transformation S_1 can consequently be calibrated accurately. If \mathbf{r} is the position vector of the spacecraft relative to the comet in Ecliptic axes, then $-\mathbf{r}$ is the position of the comet relative to the spacecraft. This vector in camera axes is $-S_1\mathbf{r}$ and, if α and β are respectively the angular x and y camera coordinates,

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = -S\mathbf{r}/r \quad (4.102)$$

where S is the matrix formed from the first two rows of S_1 (eqn 4.101).

The computation (in a subroutine) of the camera coordinates at time t from the vector of 12 parameters proceeds as follows.

1. Calculate the spacecraft position and velocity relative to the Sun at zero time from the orbital elements.
2. Form the position and velocity of the comet relative to the Sun at zero time, ie. using the position and velocity of the spacecraft relative to the comet from the parameter vector.
3. Using Kepler's equation, determine the true anomaly θ of the spacecraft orbit at time t , and hence obtain the spacecraft position and velocity (relative to the Sun) at time t .
4. Calculate the orbital elements of the comet from the position and velocity at zero time.
5. Using Kepler's equation, determine the true anomaly θ of the comet orbit at time t , and hence obtain the comet position and velocity (relative to the Sun) at time t .
6. The spacecraft position at time t (relative to the comet) is now available from eqn 4.95 and similarly for velocity.
7. The camera angular coordinates can now be obtained by means of eqn 4.102.

This function subroutine is needed to form the fitting residuals of eqn 4.42 and the matrix H of partial derivatives in eqns 4.41 and 4.43.

The example refers to rendezvous in the year 2011 near aphelion of the comet Wirtanen since this was a candidate for the *Rosetta* mission (Ref. [26]). The orbital elements of this comet have been determined by Muraoka from 83 observations between 1985 and 1997 as follows. Assume conventional notation.

a	3.0991080 AU
e	0.6567522
Time at perihelion	2450521.7 julian days
ω	356.342 deg
Ω	82.205 deg
inc	11.722 deg
period	5.456 years

The position and velocity of the spacecraft relative to the comet at zero time (0.00 hrs, 10 June 2011 or 2455722.5 julian days) were taken to be as follows.

position	-0.100	0.200	0.020	million km
velocity	0.210	-0.560	-0.050	km/s

which corresponds to an approach velocity of 0.600 km/s offset from the direction of the comet by 6.0 degrees to ensure an apparent motion of the comet against the star background. The above position vector (224,000 km from the comet) was made to correspond to the end of a 10 day fitting interval, which started at a separation of approximately 820,000 km. Samples were simulated at 6 hour intervals from the camera, and both range and range-rate from one of the ground tracking sites (Carnarvon, Kourou, Villafranca).

The following typical measurement errors (standard deviations) were assumed

measurement	random error	bias error	units
camera angles	6.0	6.0	microradians
range	12.0	12.0	m
range-rate	2.0	2.0	mm/s

Note that a representative camera might have a pixel size of 12 microradians and that a typical 5 km diameter comet subtends 10 microradians at 500,000 km.

The initial conditions were as listed in Table 4.7 but some explanation is necessary. Before the comet is acquired the orbits (relative to the Sun) of the spacecraft and comet will have been determined, the former by measurements of range and range-rate and the latter by ground-based optical observations. By using the software of these examples or otherwise it can be confirmed that, although the position of the comet may not be known more accurately than several hundred kilometres the velocity (and that of the spacecraft) should have been determined to an accuracy the order of 1 metre per second. It is for this reason that the initial estimation errors in Table 4.7 have been taken to be (pessimistically) 10,000 km in relative position but (realistically) 1 m/s in relative velocity. This is important because, as a result of using onboard camera sightings of the comet, *only the ratios of the components of relative position to relative velocity are well determined* but not the separate components of position and velocity.

Five re-linearizing iterations were necessary to achieve convergence with one batch of data and Table 4.8 shows the resulting representative estimates. They vary according to the randomization. The RMS residuals have been reduced correctly to approximately unity, that value arising due to normalization of the simulated measurements. It can be observed that, as a result of the camera sightings, the position of the spacecraft relative to the comet has been greatly improved, even along the separation vector between the spacecraft and comet. By inspection of the ratios "err/sig", we observe that the errors of this simulated estimation are within the desired range ± 2 , ie. they are consistent with the theoretical standard deviations corrected for bias.

Table 4.7. SRIF fit for Approach to a Comet: First Estimates

Parameter	estimate	error	error/sig ⁽¹⁾
e_x	0.63000	-0.00321403	-0.321
e_y	-0.02000	0.00211232	0.211
θ^* deg	188.00	0.51000	0.510
p AU	1.87	0.01316	0.394
inc deg	11.00	-0.66000	-0.660
Ω deg	82.00	-0.16000	-0.160
rel x km	-110000	-10000	-1.000
rel y km	210000	10000	1.000
rel z km	10000	-10000	-1.000
rel \dot{x} km/s	0.211	0.00100	1.000
rel \dot{y} km/s	-0.561	-0.00100	-1.000
rel \dot{z} km/s	-0.049	0.00100	1.000

¹ "sig" refers to the standard deviations of the first estimates.

Table 4.8. SRIF fit for Approach to a Comet:
Final Estimates using range, range-rate, and camera angles

Parameter	estimate	error	error/sig ⁽¹⁾
e_x	0.63323	0.00001194	0.471
e_y	-0.02209	0.00002468	0.491
θ^* deg	187.49	0.00198	0.495
p AU	1.86	-0.00006	-0.469
inc deg	11.66	-0.00311	-0.478
Ω deg	82.16	-0.00204	-0.496
rel x km	-100185.0	-185.0	-1.098
rel y km	200367.1	367.1	1.089
rel z km	20035.5	35.5	1.053
rel \dot{x} km/s	0.210	0.00045	1.277
rel \dot{y} km/s	-0.561	-0.00117	-1.253
rel \dot{z} km/s	-0.050	-0.00010	-1.244

¹ "sig" is the theoretical standard deviation corrected for bias terms.

4.5 Software

The simulations of this section were carried out by means of C++ programs grouped together in the files OBS.CPP and SUB_OBS.CPP which have been deposited in the Internet FTP server of the publisher as specified in the preface.

All the routines are in double precision; those included in the file OBS.CPP are as follows.

Main : main program (1) to generate first estimates of orbits by the methods of section 4.2, (2) to refine orbital estimates by means of SRIF in the batch mode (sections 4.3.3 and 4.3.4).

measure_srif : to transform from the parameter vector to the measurement vector.

camera : to transform from the parameter vector to onboard camera angles.

Other subroutines are grouped in the file SUB_OBS.CPP .

observe : given position and velocity vectors, calculate the ground-based angles, range, and range-rate.

reconstruct : given ground based range, azimuth, and elevation, recover the position vector of the space body.

Earth_Sun : Ephemeris of the Sun; see Appendix A.4.

r_rho : solution of equations 4.22 and 4.17; see section 4.2.2.

fn : minor function used in routine r_rho.

Laplace_first : first estimate of orbital elements according to section 4.2.2.

velocity_derived : calculation of velocity vector from position fixes as in section 4.2.4.

random_normal : normally distributed random numbers.

x_to_elements : change from the parameter vector to orbital elements.

house : matrix triangularization routine.

triang : inversion of upper triangular matrix.

Orb_rv : calculation of the position and velocity vectors from the orbital elements ($e_x, e_y, \theta^*, p, i, \Omega$).

Transform : rotational transformations; see section 3.10.

The files ORBIT2.CPP (section 1.9) and MATRIX2.CPP of orbit and matrix routines in double precision are also required.

CHAPTER 5

MIDCOURSE AND GRAVITY ASSIST MANOEUVRES

5.1 Introduction

After completion of the launch phase a spacecraft will be either in an Earth satellite orbit, or in a hyperbolic orbit that will result in escape from the Earth's sphere of influence prior to a deep space mission. We are concerned in this chapter only with the latter type of mission.

The hyperbolic escape orbit will have arisen from additional increments of velocity generated by the third and usually fourth rocket stages, following either a circular or eccentric parking orbit. At burn-out of the last stage there will be departures from the nominal trajectory, principally in the magnitude and direction of the velocity vector due to errors in pointing the rocket motors and in off-nominal performance (effective specific impulse). These errors are detected by determining the subsequent orbit by means of ground-based range and doppler (range-rate) measurements. They can then be corrected by the *first midcourse manoeuvre*, preferably as soon as possible in order to limit the required magnitude of the manoeuvre and hence conserve propellant of the onboard propulsion system. The latter would typically be a mono-propellant liquid system but a bi-propellant system (with a higher specific impulse) might be employed, especially if significant subsequent non-corrective manoeuvres are required, eg. rendezvous and orbiting of a planetary body. The first midcourse correction will not be perfect because both the orbit determination and execution of the manoeuvre will be subject to errors. *Subsequent midcourse manoeuvres* will consequently usually be necessary as analyzed below.

The performance required from the three or four stages of the launch system to achieve a 500-1000 kg spacecraft for planetary missions is demanding. Therefore ways of enhancing performance have been sought, the most important being (a) gravity assist manoeuvres as explained and illustrated in this chapter, and (b) electric propulsion, the subject of the next chapter.

5.2 Midcourse Manoeuvres

5.2.1 Nominal Trajectory to Jupiter

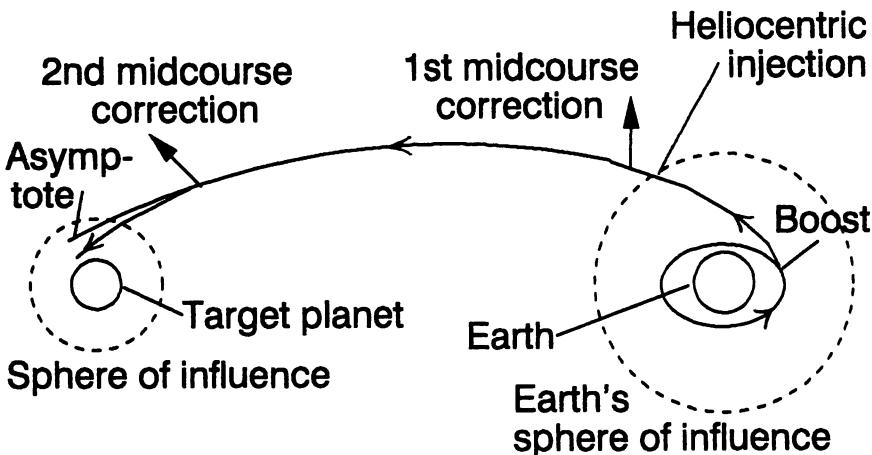


Fig. 5.1. Midcourse Corrections

In order to introduce illustrative computations this section (on midcourse corrections) is developed against the background of the representative space mission of Figure 5.1. It is the first part (to Jupiter) of the Solar Polar Mission of section 1.7.3. The space vehicle is initially in a circular parking orbit at an altitude of 185 km, the orbital elements (in equatorial axes) being as follows.

a	6563.0 km
e	0.000
θ^*	-16.1 deg at start time
start time	3.96 hrs 8 Oct 1990, 2448172.67 julian days
Ω	10.1 deg
inc	29.0 deg

The spacecraft velocity is then boosted impulsively by 7.580 km/s along the tangent of the circular orbit. This is an idealised calculation because the velocity increment would not be an impulse and because two stages might be necessary, but the approximation will suffice for this example. As a result, the subsequent orbit is a hyperbola which corresponds to eventual escape from the gravitational sphere of influence of the Earth at a radius of 0.93 million km (Table 1.2), with $v(\infty)$ equal to 10.718 km/s. After adding the Earth's velocity around the Sun, the orbital elements (in Sun-centred ecliptic axes) are as follows.

a	6.598 AU
aphelion	12.2 AU
e	0.848
θ	2.6 deg at start time
start time	3.96 hrs 8 Oct 1990, 2448172.67 julian days
ω	-1.3 deg
Ω	15.3 deg
inc	1.6 deg

and the position and velocity (in Sun-centred ecliptic axes) at heliocentric injection are

AU	km/s
0.959	-10.739
0.286	39.018
0.001	1.150

As a result of adjusting (a) the duration of the interplanetary transfer, (b) the launch date, and (c) $(\theta^* - \Omega)$ of the parking orbit for the required inclination, the nominal position and velocity (relative to Jupiter in ecliptic axes) 516.4 days after boost are

AU	km/s
-0.0001	-8.106
0.0002	9.436
-0.0002	-0.120

5.2.2 Errors at Heliocentric Injection

The principal errors at heliocentric injection arise from the velocity vector increment generated by the last stage, ie. in boosting from the parking orbit. If \mathbf{r}_0 and \mathbf{v}_0 are the position and velocity in the circular parking orbit at boost, the normal to that orbit is

$$\mathbf{i}_z = \text{unit}(\mathbf{r}_0 \times \mathbf{v}_0) \quad (5.1)$$

Denote the unit vectors of position and velocity as $\tilde{\mathbf{r}}_0$ and $\tilde{\mathbf{v}}_0$. The boost velocity errors are taken to be (a) a proportional error ϵ_1 in the magnitude, and (b) angular errors ϵ_2 and ϵ_3 in direction. Thus, if the nominal boost velocity vector is $\Delta\mathbf{v}$, the actual velocity vector after boost is

$$\mathbf{v}_1 = \Delta\mathbf{v}[(1 + \epsilon_1)\tilde{\mathbf{v}}_0 + \epsilon_2\tilde{\mathbf{r}}_0 + \epsilon_3\mathbf{i}_z] + \mathbf{v}_0 \quad (5.2)$$

The position and velocity can be propagated forward from boost to heliocentric injection (radius of 0.93 million km) by the methods of sections 1.3.2 and 1.3.1, eg. the computer subroutines of section 1.9. Such a computation can be repeated six times with small perturbations in the three components of position and velocity at boost to generate numerically the 36 partial derivatives between position and velocity at boost and heliocentric injection, ie. the *state transition matrix* denoted Φ . It is convenient now to introduce some notation for sub-matrices as follows.

$$\Phi = \begin{bmatrix} \Phi_a & \Phi_b \\ \Phi_c & \Phi_d \end{bmatrix} \quad (5.3)$$

$$\Lambda = \begin{bmatrix} \Phi_b \\ \Phi_d \end{bmatrix} \quad (5.4)$$

$$\Psi = [\Phi_a \Phi_b] \quad (5.5)$$

If δx denotes the six components of the state vector (position and velocity) at heliocentric injection and δv_1 is the vector of boost velocity errors, then

$$\delta x = \Lambda \delta v_1 \quad (5.6)$$

Therefore

$$\delta x \delta x^T = \Lambda \delta v_1 \delta v_1^T \Lambda^T \quad (5.7)$$

Take expected values to yield the covariance matrix of errors at heliocentric injection

$$G_1 = \Lambda G_0 \Lambda^T \quad (5.8)$$

where G_0 is the 3x3 covariance matrix of errors in the velocity vector v_1 after boost, eg. a diagonal matrix of the variances corresponding to the errors in eqn 5.2. If the errors in the boost velocity vector increment are taken to be 1 per cent in magnitude and 1 degree in pointing, then the covariance matrix at heliocentric injection G_1 is computed such that the root-sum-squares position and velocity errors (sum of three diagonal terms) are 16553 km and 194 m/s respectively. The latter errors would give rise to enormous dispersion at Jupiter and consequently it is necessary to remedy them with a midcourse correction. This manoeuvre should be fairly early because the required magnitude of the correction increases if it is delayed significantly. On the other hand enough time must have elapsed to permit a sufficiently accurate determination of the heliocentric orbit. These computations have been carried out as if the first midcourse manoeuvre is made at heliocentric injection which, at approximately 24 hours after boost from the parking orbit, is typical.

5.2.3 Fixed Time Guidance

Let now Φ and the sub-matrices of eqn 5.3 refer to the 515 day interval from heliocentric injection at time t_1 to the encounter with Jupiter at time t_f . Such a transition matrix is computed as before by means of numerically perturbed trajectories. If the vector of position errors to be corrected at the final time is

$$\delta r(t_f) = \Psi \delta x(t_1) \quad (5.9)$$

then the required manoeuvre at time t_1 satisfies

$$-\delta r(t_f) = \Phi_b \Delta v_1(t_1) \quad (5.10)$$

hence

$$\Delta v_1(t_f) = -\Phi_b^{-1} \Psi \delta x(t_1) \quad (5.11)$$

The mean-squared value of this velocity change is calculated by post multiplying the above by Δv_1^T and taking expected values.

Thus

$$\bar{\Delta v}_1^2 = \Phi_b^{-1} P_0 \Phi_b^{-T} \quad (5.12)$$

where (using eqn 5.8)

$$P_0 = \Psi G_1 \Psi^T \quad (5.13)$$

The above strategy is described as *fixed time guidance* because all three components of position are zeroed at the nominal fixed final time t_f .

Such a manoeuvre is not however executed perfectly and furthermore, the orbit determination on which it is based is also imperfect. Let us calculate the errors arising from (a) an execution error $\delta \hat{v}_1$, and (b) orbit determination errors $\delta \tilde{x}(t_1)$. The resulting position error vector at the final time is

$$\delta r(t_f) = \Psi \delta \tilde{x}(t_1) + \Phi_b \delta \hat{v}_1 \quad (5.14)$$

Post multiply by the transpose and take expected values, to give the 3x3 covariance matrix of position errors at time t_f , viz.

$$P_1 = \Psi \tilde{G}_1 \Psi^T + \Phi_b \hat{G}_1 \Phi_b^T \quad (5.15)$$

The 6x6 matrix \tilde{G}_1 comes from the orbit determination and \hat{G}_1 is the 3x3 matrix of execution errors, ie. in controlling the magnitude and pointing of the onboard propulsion system.

5.2.4 Variable Time Guidance

The dispersion represented by eqn 5.15 will be a big improvement but will usually still be unacceptably large. A second or even third midcourse correction is therefore common. Let this occur at time t_2 ; Φ now refers to the time interval t_2 to t_f . In calculating this second manoeuvre it is postulated that it is *variable time guidance*; an encounter occurs but we relax the requirement for it to be exactly at the nominal time t_f . The choice of one fixed time manoeuvre and a second variable time adjustment is arbitrary but sensible, although the choice is largely for illustrative purposes.

To be more specific, let \mathbf{v}_{inf} be the approach velocity vector relative to Jupiter, ie. along the asymptote of the hyperbola within the sphere of influence of Jupiter. Define a set of axes RST centred at Jupiter as follows. The unit vector \mathbf{e}_S is along the S axis and is given by

$$\mathbf{e}_S = \text{unit}(\mathbf{v}_{inf}) \quad (5.16)$$

The unit vector \mathbf{e}_T is perpendicular to \mathbf{e}_S and the ecliptic plane (with normal \mathbf{k}). Therefore

$$\mathbf{e}_T = \text{unit}(\mathbf{e}_S \times \mathbf{k}) \quad (5.17)$$

and finally

$$\mathbf{e}_R = \mathbf{e}_S \times \mathbf{e}_T \quad (5.18)$$

If these vectors are in ecliptic axes the transformation from RST to ecliptic axes is

$$[\mathbf{e}_R \quad \mathbf{e}_S \quad \mathbf{e}_T] \quad (5.19)$$

Variable time guidance arises from transforming the final dispersion vector δr into RST axes and ignoring errors along the S axis, since these (at least for small errors) are associated only with variations in the arrival time.

If x and y are the miss components along the T and R axes

$$\begin{bmatrix} x \\ y \end{bmatrix} = \Gamma \delta r(t_f) \quad (5.20)$$

where the 2×3 matrix Γ is formed from the third and first rows of the transpose of eqn 5.19. Thus the necessary corrective manoeuvre is given by

$$-\begin{bmatrix} x \\ y \end{bmatrix} = \Gamma \Phi_b \Delta v_2 \quad (5.21)$$

but there is a degree of freedom because there are three components of velocity to satisfy only two equations. The degree of freedom can be used to minimize the squared magnitude of Δv_2 while satisfying the two equality constraints of eqn 5.21. The details of introducing two Lagrangian multipliers are left to the student, the solution being

$$\Delta v_2 = -A^T [AA^T]^{-1} \begin{bmatrix} x \\ y \end{bmatrix} \quad (5.22)$$

where

$$A = \Gamma \Phi_b \quad (5.23)$$

The final dispersion due to orbit determination and execution errors at time t_2 is calculated as at eqn 5.14 except the $\delta r(t_f)$ of that equation is transformed into only the components in the RT plane. Thus

$$\begin{bmatrix} x \\ y \end{bmatrix} = \Gamma [\Psi \delta \tilde{x}(t_2) + \Phi_b \delta \hat{v}_2] \quad (5.24)$$

where Ψ and Φ_b are now sub-matrices of the transition matrix Φ from t_2 to t_f . By post multiplying and taking expected values we obtain the 2×2 covariance matrix of the miss- components as

$$Q_2 = \Gamma P_2 \Gamma^T \quad (5.25)$$

where

$$P_2 = \Psi \tilde{G}_2 \Psi^T + \Phi_b \hat{G}_2 \Phi_b^T \quad (5.26)$$

\tilde{G}_2 and \hat{G}_2 are respectively the covariance matrices of the orbit determination and execution errors at the second midcourse manoeuvre at time t_2 .

If the errors can be approximated as normal then, as shown in any standard text on probability, the probability density of the final dispersion is given by

$$\frac{1}{2\pi|Q_2|^{1/2}} \exp\left(-(1/2) [x \ y] Q_2^{-1} [x \ y]^T\right) \quad (5.27)$$

The contours for a given probability level are ellipses. The semi-axes are calculated from the square roots of the two solutions (eigenvalues of Q_2)

$$s(\pm) = (1/2) \left[(q_{11} + q_{22}) \pm ((q_{11} - q_{22})^2 + 4q_{12}^2)^{1/2} \right] \quad (5.28)$$

and the major semi-axis is at angle α to the x or T axis, where

$$\tan \alpha = [s(+)-q_{11}]/q_{12} = q_{12}/[s(+) - q_{22}] \quad (5.29)$$

For multiples of n of the above semi-axes, the probability level P is

$$P = 1 - \exp(-n^2/2) \quad (5.30)$$

e.g. when $n=2.4475$, $P=95$ per cent.

5.2.5 Parametric Results

Some illustrative results are included for different sets of some of the parameters. They all refer to the nominal trajectory to Jupiter already introduced. The baseline choice of errors (standard deviations) is as follows.

Boost velocity errors : 1.0 per cent in magnitude and 1 degree in pointing.
Manoeuvre velocity errors : 1.0 per cent in each component.

Orbit determination at 1st manoeuvre : 100 km in each component of position and 0.1 m/s in each component of velocity.

Orbit determination at 2nd manoeuvre : 200 km in each component of position and 0.1 m/s in each component of velocity.

However, as a simplification, the covariance matrices of manoeuvre and orbit determination errors have been taken to be diagonal. The full covariance matrix of the latter would be a side product of the actual orbit determination. Further details of the baseline case are as follows and parametric variations are listed in Table 5.1.

1. RSS position error at Jupiter if no midcourse corrections are made = 0.09851 AU.
2. First correction is made 23.72 hours after boost employing fixed time guidance. RMS magnitude = 193.9 m/s.
3. RSS position error at Jupiter after first midcourse correction = 0.00156 AU.

Table 5.1. Midcourse Guidance to Jupiter

Orbit det. errors	t_2 days	Execution ⁽²⁾ errors	Final RSS miss km	Δv_1 m/s	Δv_2 m/s	$(\Delta v_1 + \Delta v_2)$ m/s
base ⁽¹⁾	300	base	3006	193.9	5.8	199.7
base	400	base	2088	193.9	10.7	204.6
base	450	base	1736	193.9	18.6	212.5
base	483	base	1589	193.9	37.0	230.9
base x 3	450	base	2993	193.9	18.8	212.7
base / 3	450	base	1534	193.9	18.6	212.5
base	450	base x 3	40690	581.4	167.3	748.7
base	450	base / 3	860	69.4	2.3	66.7

¹ base is shorthand for baseline parameter.

² including the boost velocity increment.

4. RSS two-dimensional error (RT plane) at Jupiter after first midcourse correction = 0.00071 AU.
5. Second correction is made at 400.0 days after boost using variable time guidance. RMS magnitude = 10.7 m/s.
6. RSS two-dimensional error (RT plane) at Jupiter after second midcourse correction = 2088 km.

Table 5.1 illustrates (1) propellant usage is dominated by the first manoeuvre, (2) a small final dispersion requires a late final correction, and (3) manoeuvre execution errors (not orbit determination) are the major influence on final dispersion.

5.3 Gravity Assist Manoeuvres

5.3.1 Introduction

Let us first examine the propulsion requirements to inject a spacecraft (initially in a circular orbit about the Earth) into a heliocentric orbit with an aphelion equal to that of a planet, ie. the minimum energy to encounter that planet. The semi-axes of the planets are listed in Table 1.1 and the required hyperbolic escape velocity v_∞ is equivalent to the Δv_1 of eqn 1.94. The velocity increment Δv to be applied from the circular Earth orbit to achieve a given v_∞ is given at eqn 1.105. The payload ratio (useful mass divided by initial mass) is quoted at eqn 2.25 as a function of the structural efficiency, specific impulse, and required velocity increment. All of these formulae have been utilized to prepare Table 5.2 which shows the payload ratios for one and two stages of propulsion applied in the circular Earth orbit. Representative values of the parameters have been assumed as follows:

1. altitude of circular Earth orbit is 185 km.

2. specific impulse of a bi-propellant system is 300 seconds.
3. structural efficiency (mass of structure divided by mass of structure and propellant) is 10 per cent.

Table 5.2. Minimum Propulsion Requirements to encounter the Planets

Planet	Semi-axis AU	v_∞ km/s	Δv km/s	Payload one stage	ratio per cent two stages
Mercury	0.387	-7.533	5.556	5.71	10.32
Venus	0.723	-2.495	3.507	22.6	25.1
Mars	1.524	2.945	3.615	21.4	24.0
Jupiter	5.203	8.793	6.306	1.93	7.26
Saturn	9.539	10.289	7.284	0.00	4.46
Uranus	19.18	11.280	7.978	0.00	3.08
Neptune	30.06	11.654	8.247	0.00	2.64
Pluto	39.44	11.813	8.363	0.00	2.47

Table 5.2 illustrates how conventional chemical propulsion permits missions to (a) Mars, Venus, and possibly Mercury with one stage from a parking orbit, (b) otherwise only Jupiter might just be feasible with two stages. In other words, *exploration of the outer planets is difficult in terms of conventional chemical propulsion*.

It was realised at an early stage that when a spacecraft passes very close to a planetary body it can pick up energy from that body, depending on the geometry and relative motion. For example, Ehricke (Ref. [18]) gave one of the first analyses of *lunar swingby manoeuvres* to enhance trajectories to Venus, Mars, and Jupiter. The famous *Grand Tour* of Voyager 2 starting in August 1977 included all the outer planets (except Pluto) culminating in a flyby of Neptune in August 1989; it would not have been possible without the energy-stealing swingby manoeuvres. The International-Sun-Earth-Explorer was re-programmed after completing its original mission, to exploit lunar swingbys and approach the comet Giacobini-Zinner in 1985. The use of swingbys of the Moon, Mars, Venus, or the Earth itself have consequently become more and more common in planning high energy planetary missions, ie. to circumvent the limitations of heliocentric injection with chemical propulsion (Refs. [15], [53], [59], [54]).

5.3.2 Gravity Assist in Two Dimensions

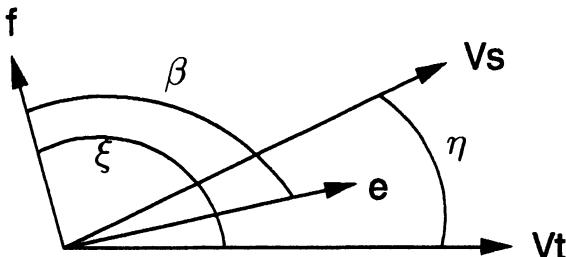


Fig. 5.2. Gravity Assist in Two Dimensions

The swingby gravity-assist manoeuvre is first analyzed in two dimensions because it is commonly applied in this way and it is easier to understand (before generalising to three dimensions). The elements of Figure 5.2 are defined as follows:

1. \mathbf{V}_s is the initial velocity vector of the spacecraft in Sun-centred ecliptic axes (heliocentric axes).
2. \mathbf{V}_t is the corresponding velocity vector of the target planetary body.
3. \mathbf{f} is the unit vector of the asymptote along which the spacecraft appears to approach to an observer on the target body.
4. \mathbf{e} is the unit vector of the asymptote along which the spacecraft appears to depart (to an observer on the target body) after a swingby.
5. angles β , ξ , and η are as defined in the figure.

The initial velocity vector of the spacecraft relative to the target body is

$$\Delta \mathbf{v} = \mathbf{f} \Delta v = \mathbf{V}_s - \mathbf{V}_t \quad (5.31)$$

Therefore the final velocity vector after the swingby is

$$\mathbf{V}_f = \mathbf{V}_t + \mathbf{e} \Delta v \quad (5.32)$$

In terms of the angle η , where \mathbf{i} and \mathbf{j} are unit vectors along the x and y axes,

$$\Delta \mathbf{v} = \mathbf{i}(V_s \cos \eta - V_t) + \mathbf{j} V_s \sin \eta \quad (5.33)$$

Therefore the angle ξ of Figure 5.2 is given by

$$\xi = \arctan \frac{V_s \sin \eta}{V_s \cos \eta - V_t} \quad (5.34)$$

and

$$\Delta v^2 = (V_s \cos \eta - V_t)^2 + V_s^2 \sin^2 \eta + 2V_s \sin \eta (V_s \cos \eta - V_t) \quad (5.35)$$

The angle through which the approach velocity vector is turned (as seen by an observer on the target planet) is obtained from eqn 1.61 for the asymptotes as

$$\beta = 2 \arcsin \left[\frac{1}{1 + \Delta v^2 r_{min} / \mu} \right] \quad (5.36)$$

where r_{min} is the radius of closest approach to the target body. Depending on which side of the target is passed, this angle could however be in the reverse sense to that illustrated in Figure 5.2. The other solution is ignored here because it would lead to a swingby with energy loss. From the figure

$$\mathbf{e} \cdot \mathbf{V}_t = V_t \cos(\xi - \beta) \quad (5.37)$$

hence permitting from eqn 5.32

$$V_f = \sqrt{V_t^2 + \Delta v^2 + 2\Delta v V_t \cos(\xi - \beta)} \quad (5.38)$$

The angle through which the spacecraft approach velocity vector is deflected is not β ; it is the smaller angle δ given by

$$\delta = \arccos \left[\frac{\mathbf{V}_s \cdot \mathbf{V}_f}{V_s V_f} \right] \quad (5.39)$$

where

$$\mathbf{V}_s \cdot \mathbf{V}_t = V_s V_t \cos \eta + \Delta v V_s \cos(\xi - \beta - \eta) \quad (5.40)$$

In summary, given the inputs (V_s, V_t, η, r_{min}), the outputs (V_f, δ) are determined.

Example 5.3.1. On 5 March 1979 Voyager 1 encountered Jupiter, the input parameters being as follows (Ref. [9]).

V_s	12.60 km/s
V_t	12.83 km/s
η	50.1 deg
r_{min}	3.484×10^5 km

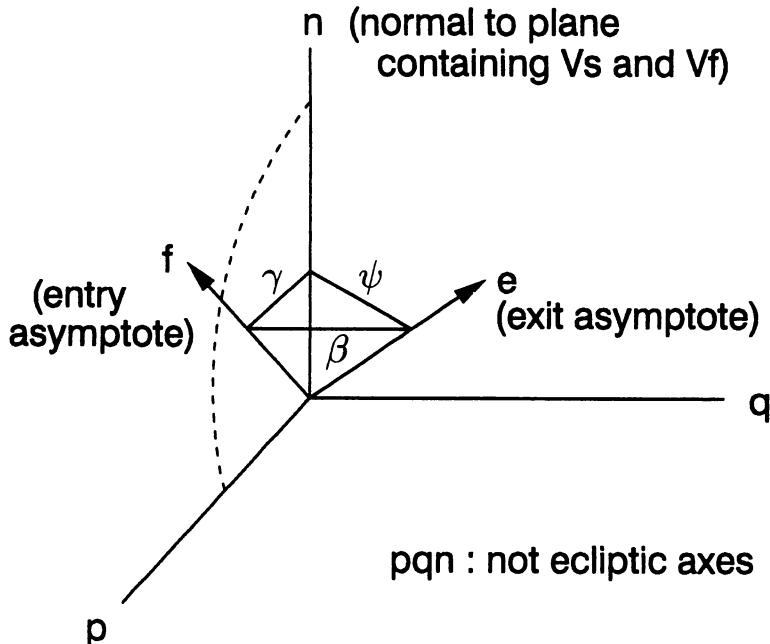
The closest approach to Jupiter (r_{min}) was 4.9 radii resulting in

V_f	23.33 km/s
β	98.6 deg
δ	58.0 deg

Thus, as a result of the swingby, the heliocentric spacecraft velocity vector has been increased by the factor 1.85 in magnitude and changed 58.0 degrees in direction.

5.3.3 Gravity Assist in Three Dimensions

With reference to Figure 5.3 the above analysis is now extended to three dimensions. Let a unit normal \mathbf{n} be defined that contains the input and output spacecraft velocity vectors, viz.

**Fig. 5.3.** Gravity Assist in Three Dimensions

$$\mathbf{n} = \text{unit}(\mathbf{V}_s \times \mathbf{V}_f) \quad (5.41)$$

It is possible to change the orbital plane of the spacecraft by the swingby, depending on how the spacecraft is guided in the approach to the target body, eg. above, below, or to the side. Therefore the unit vector \mathbf{n} is assumed to be a specified parameter. The definition of the pqr axes of Figure 5.3 is completed by requiring that the entry asymptote \mathbf{f} is in the pn plane. The angle β is as defined above at eqn 5.36.

To calculate the angle ψ of Figure 5.3 take the scalar product of eqn 5.32 with \mathbf{n}

$$\mathbf{V}_f \cdot \mathbf{n} = 0 = \mathbf{V}_t \cdot \mathbf{n} + \Delta v \mathbf{e} \cdot \mathbf{n} \quad (5.42)$$

Since $\mathbf{e} \cdot \mathbf{n} = \cos \psi$ it follows that

$$\psi = \arccos \left[-\frac{\mathbf{V}_t \cdot \mathbf{n}}{\Delta v} \right] \quad (5.43)$$

where the magnitude Δv is calculated from eqn 5.31. The angle γ of Figure 5.3 is obtained from

$$\gamma = \arccos(\mathbf{n} \cdot \mathbf{f}) \quad (5.44)$$

Let the exit asymptote in pqr axes be

$$\mathbf{e} = (\sin \psi \cos \phi, \sin \psi \sin \phi, \cos \psi) \quad (5.45)$$

and note that

$$\cos \beta = \mathbf{e} \cdot \mathbf{f} = \sin \psi \cos \phi \sin \gamma + \cos \psi \cos \gamma \quad (5.46)$$

Therefore

$$\phi = \pm \arccos \left[\frac{\cos \beta - \cos \psi \cos \gamma}{\sin \psi \sin \gamma} \right] \quad (5.47)$$

The plus or minus sign appears in eqn 5.47 as a reminder to check the two possible swingby manoeuvres (for a given β) referred to previously. Only one case is illustrated in the figure.

The exit asymptote \mathbf{e} in eqn 5.45 is given in pqn axes and it is required in the ecliptic axes assumed to be employed to specify \mathbf{V}_t, \mathbf{n} , etc. Note that \mathbf{q} in ecliptic axes is

$$\mathbf{q} = \text{unit}(\mathbf{n} \times \mathbf{f}) \quad (5.48)$$

and

$$\mathbf{p} = \mathbf{q} \times \mathbf{n} \quad (5.49)$$

It follows that the transformation from pqn to ecliptic axes is the orthogonal matrix

$$[\mathbf{p} \quad \mathbf{q} \quad \mathbf{n}] \quad (5.50)$$

This completes the calculation of the intermediate variables necessary to calculate the spacecraft final velocity vector \mathbf{V}_f (eqn 5.32), given the inputs $(\mathbf{V}_t, \mathbf{V}_s, r_{min}, \mathbf{n})$. Finally the deflection of the spacecraft velocity vector is (as before) given by eqn 5.39.

5.3.4 An Example : Comet Nucleus Sample Return

As already mentioned, the use of one or more gravity assist manoeuvres has become almost standard practice in the design of new interplanetary missions. Such exercises involve the use of Lambert's algorithm and patched conics to link several different orbits both with respect to the Sun and planets subject to swingby manoeuvres. All the required calculations have been included in chapter 1; see also section 1.9 for a list of the applicable software. At a basic level, computer programs for such mission design have been interactive trial-and-error procedures but, for more sophisticated studies, they have been integrated into schemes of constrained optimization to permit minimization of propellant usage. The search for a feasible sequence of orbits and swingbys for long term missions (especially to the outer planets) has been automated by JPL (Ref. [59]).

As an example, details are given here of the use of supplied illustrative software which exploits subroutines already introduced (section 1.9) for patched conics and Lambert's algorithm (section 1.5). The mission is *Comet Nucleus Sample Return* (CNSR) as originally studied by the European Space

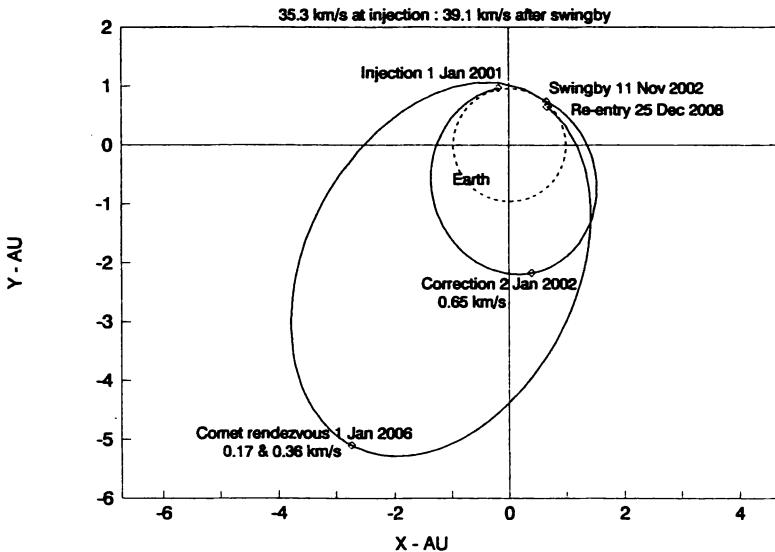


Fig. 5.4. Gravity Assist for a Comet Mission

Agency in the 1980's (Ref. [2]). It was to orbit and land on a comet nucleus, and to return to Earth with samples from the comet. At the time of writing, the project is proceeding (Ref. [5]) under the name of ROSETTA but without the return to Earth. A feasible representative trajectory is shown in Figure 5.4 comprising the following segments (orbits).

1. From heliocentric injection to approximately the aphelion (2.21 AU) of an intermediate orbit where a small correction is applied at 366 days.
2. From this correction to a close approach to the Earth (300 km altitude) for a swingby manoeuvre at 679 days.
3. From the swingby to rendezvous with the comet at 1826 days, approximately at aphelion (5.79 AU) of the larger orbit and that of the comet. Manoeuvres to rendezvous with and depart from the comet such as those shown on Figure 5.4 are necessary.
4. Re-entry into the atmosphere of the Earth at 2915 days, ie. approximately 8 years from the start.

There is some scope for minimizing propellant requirements, subject to the constraint of comet rendezvous, by adjusting the times of the correction, swingby, and re-entry. Further numerical details are as follows.

Injection velocity v_∞	5.05 km/s
Total velocity after injection	35.34 km/s
Hyperbolic deflection β	47.1 deg
Deflection of heliocentric velocity vector δ	9.7 deg
Total velocity after swingby	39.11 km/s
Correction at 366 days to adjust swingby	650 m/s

Thus, at a cost of a longer mission, use of Earth-gravity-assist yields effectively a crucial extra 3.77 km/s at the delayed "injection" after swingby.

5.4 Software

The simulations of this section were carried out by means of C++ programs grouped together in the files MID.CPP (for midcourse guidance) and GRAV.CPP (for gravity assist manoeuvres). All the routines are in double precision and linking to the previously introduced subroutines in ORBIT2.CPP and MATRIX2.CPP is necessary.

The routines included in the file MID.CPP are as follows.

Main : main program to illustrate two midcourse manoeuvres to Jupiter as presented in section 5.2.

Earth_RV : position and velocity of Earth given Julian days.

Jupiter_RV : position and velocity of Jupiter given Julian days.

propagate : see section 3.10.

sub_STM : to select sub-matrices of the state transition matrix.

covariance : to form the matrix $PQPT^T$.

diag_sum : square root of the sum of the diagonal elements of a matrix.

gam_v : from the approach asymptote vector to RST axes (eqn 5.19).

The routines included in the file GRAV.CPP are as follows.

Main : main program to illustrate calculation of the Earth-gravity-assist manoeuvre and the patched conics for a mission to rendezvous with a comet, as presented in section 5.3.

iomeg : (i, Ω, ω) calculated from the normal to the orbit plane and a reference direction.

swingy : from $(\mathbf{V}_t, \Delta\mathbf{v}, \mathbf{n}, \beta)$ to $\mathbf{e}\Delta\mathbf{v}$.

CHAPTER 6

LOW THRUST MISSIONS

6.1 Introduction

The limitations of conventional chemical propulsion for missions to the outer planets, comets, and asteroids were illustrated in Table 5.2 of the previous chapter. An alternative approach to the use of swingby manoeuvres is to exploit the very high specific impulses that can be achieved through the use of electric propulsion, eg. exhaust velocities the order of 30 km/s instead of about 3 km/s associated with chemical propellants. The source of the electrical energy can be either solar (solar electric propulsion SEP) or nuclear (nuclear electric propulsion NEP). As a result of the diminishing power of the Sun with distance, SEP is not useful beyond about 5 AU from the Sun. Nuclear energy generation is of course independent of the Sun but has unfortunately become less attractive in recent years due to an influential Earth-bound environmental lobby.

6.2 Electric Propulsion

Thrust is generated in an electrical propulsion system by accelerating an electrically charged exhaust (usually an ionised gas) by means of electrostatic or electromagnetic forces. There is no fundamental limit to the resulting exhaust velocity, other than the velocity of light, although the required power unit represents a practical limitation. The efficiency η of an electric propulsion unit can be expressed as

$$\eta = \frac{m_f c_e^2}{2P} \quad (6.1)$$

m_f is the mass flow rate, c_e is the exhaust velocity, and P is the input power. Since thrust $F = m_f v$ and specific impulse I_{sp} is given by $F/m_f g$, it follows that the thrust developed can be calculated from

$$F = \frac{2\eta P}{I_{sp}g} \quad (6.2)$$

There are three principal categories of electrical propulsion systems as follows.

Electrostatic Propulsion : Ion Engines, Stationary Plasma Thrusters, Field Emission Thrusters.

Electrothermal Propulsion : Arcjets, Resistojets, Power Augmented Catalytic Thrusters.

Electromagnetic Propulsion : Magneto Plasma Dynamic Thrusters.

Such systems have been under development in Europe, USSR, and Japan for many years and small thrusters (order of 10 mN thrust) have been flown on satellites either for station-keeping of communication satellites or (especially the USSR) for experimental purposes. Electrical thrusters relevant to this chapter are defined as *primary propulsion systems* to permit high energy missions in deep space, requiring thrust levels in the range 100mN to 1N. Despite earlier work in the USA (Ref. [52]), development of systems with these higher levels of thrust is now taking place under the *aegis* of the European Space Agency (Ref. [46]). A new ion unit producing 200-300 mN thrust is being developed called the "ESA-XX" which is the outcome primarily of (a) German experience with the RIT-35 engine using either Mercury, Argon, or Xenon, and (b) British research culminating in the UK-25 Xenon thruster. Both solar electric and nuclear electric propulsion (Ref. [35]) using possibly the Russian Topaz nuclear power unit are possible.

6.3 Optimization of Low Thrust Trajectories

When thrust is applied continuously over an extended period of time in order to achieve the required final conditions of a space mission, it is by no means obvious how the thrust should most efficiently be applied. Should the level of thrust be varied and in what direction should the thrust vector be pointed? To analyze this problem we need to apply the classical calculus of variations or equivalently the more modern optimal control theory. The latter is treated in many standard texts such as Bryson and Ho (Ref. [8]) or Noton (Ref. [40]). It is not appropriate here to derive the basic theoretical results so, in the interests of readers not familiar with this theory, the treatment will be kept as brief and simple as possible.

6.3.1 Optimization in Two Dimensions

Let us first consider the problem in two dimensions of transferring from given initial conditions to specified final conditions (the terminal constraints) while maximizing the final mass of the space vehicle. The time for the transition is fixed as from time t_0 to t_f . In polar coordinates the equations of motion are

$$\begin{aligned}\ddot{r} &= f_r - \mu/r^2 \\ \ddot{\theta} &= (f_\theta - 2\dot{r}\dot{\theta})/r \\ \dot{m} &= -cu_1\end{aligned}\tag{6.3}$$

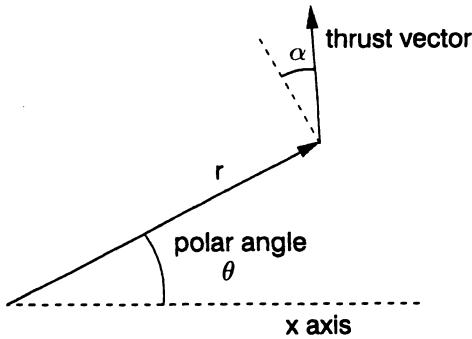


Fig. 6.1. Low Thrust in Two Dimensions

where m is the vehicle mass. The radial and transverse accelerations are given by

$$f_r = (bu_1/m) \sin \alpha \quad f_\theta = (bu_1/m) \cos \alpha \quad (6.4)$$

where b and c are constants. The angle α is defined in Figure 6.1. Define the five state variables as $x_1 = r, x_2 = \theta, x_3 = \dot{r}, x_4 = \dot{\theta}, x_5 = m$, hence

$$\begin{aligned} \dot{x}_1 &= x_3 \\ \dot{x}_2 &= x_4 \\ \dot{x}_3 &= f_r - \mu/x_1^2 \\ \dot{x}_4 &= (f_\theta - 2x_3x_4)/x_1 \\ \dot{x}_5 &= -cu_1 \end{aligned} \quad (6.5)$$

In this notation we wish to maximize $x_5(t_f)$ subject to the terminal constraints at time t_f

$$x_1 = r_f \quad x_2 = \theta_f \quad x_3 = 0 \quad x_4 = V_\theta/r_f \quad (6.6)$$

where terms on the right hand sides are specified. Equation 6.6 corresponds to the rendezvous problem when the spacecraft and target position and velocity must coincide at a given time. The so called Hamiltonian is (Refs. [8] or [40])

$$\begin{aligned} H &= \lambda_1 x_3 + \lambda_2 x_4 + \lambda_3 (f_r - \mu/x_1^2) \\ &+ \lambda_4 (f_\theta - 2x_3x_4)/x_1 - \lambda_5 cu_1 \end{aligned} \quad (6.7)$$

where the costate variables satisfy the following equations obtained from $\dot{\lambda}_i = -\partial H / \partial x_i$.

$$\begin{aligned} \dot{\lambda}_1 &= -2\lambda_3\mu/x_1^3 - 2\lambda_4x_3x_4/x_1^2 \\ \dot{\lambda}_2 &= 0 \\ \dot{\lambda}_3 &= \lambda_1 - 2\lambda_4x_4/x_1 \\ \dot{\lambda}_4 &= \lambda_2 - 2\lambda_4x_3/x_1 \\ \dot{\lambda}_5 &= -\lambda_3f_r/x_5 - \lambda_4f_\theta/x_5 \end{aligned} \quad (6.8)$$

The four terminal constraints of eqn 6.6 are adjoined to the function to be maximized (the final mass) by means of four Lagrangian multipliers ($\mu_i, i = 1, 2, 3, 4$). Thus the total function to be maximized becomes

$$V = -x_5 + \mu_1(x_1 - r_f) + \mu_2(x_2 - \theta_f) + \mu_3x_3 + \mu_4(x_4 - V_\theta/r_f) \quad (6.9)$$

evaluated at the terminal time, and the required terminal conditions arise from $\lambda_i(t_f) = \partial V / \partial x_i$.

$$\begin{aligned} \lambda_1(t_f) &= \mu_1 \\ \lambda_2(t_f) &= \mu_2 \\ \lambda_3(t_f) &= \mu_3 \\ \lambda_4(t_f) &= \mu_4 \\ \lambda_5(t_f) &= -1 \end{aligned} \quad (6.10)$$

For any chosen control strategy (to be optimized), the state equations 6.5 can be integrated forward from the specified initial conditions. The costate equations 6.8 must be integrated backwards from the final conditions 6.10 which depend on the unknown constants μ_i . The optimization strategy follows from Pontryagin's Maximum Principle (Refs. [8] or [40]) which states that the function H of eqn 6.7 is to be maximized at all times by choice of the control variables u_1 and the pointing angle α . The latter process depends on the costate variables so, except for few very simple problems, a numerical procedure becomes necessary in which the four multipliers ($\mu_i, i = 1, 2, 3, 4$) are adjusted iteratively to satisfy the terminal conditions 6.6.

It is not necessary here for us to carry out the numerical procedure. The *form of the optimization strategy* can be ascertained by applying the Maximum Principle to the function H of eqn 6.7. Let us first note the necessary condition for H to be a maximum with respect to the control variable α .

$$\frac{\partial H}{\partial \alpha} = (bu_1/x_5) [\lambda_3 \cos \alpha - \lambda_4 \sin \alpha] \quad (6.11)$$

i.e. provided u_1 is non-zero, the optimum pointing angle is determined by

$$\alpha = \arctan(\lambda_3/\lambda_4) \quad (6.12)$$

By substituting eqn 6.12 into 6.7 we can write H in the form

$$H = Qu_1 + R \quad (6.13)$$

where Q and R are independent of u_1 . In particular

$$Q = (b/x_5) \sqrt{\lambda_3^2 + \lambda_4^2} - c\lambda_5 \quad (6.14)$$

The optimum u_1 is not determined from eqn 6.13 because $\partial H / \partial u_1$ is independent of u_1 . However, if we assume that there is an upper limit u_{max} to the available thrust and that it is non-negative, application of the Maximum Principle leads to the following conclusions.

$$\begin{aligned} Q > 0 &\text{ implies } u_1 = u_{max} \\ Q < 0 &\text{ implies } U_1 = 0 \\ Q = 0 &\text{ implies } \partial^2 H / \partial u_1^2 = 0 \end{aligned}$$

The third possibility is an uncommon singularity which we will ignore (singular control). Otherwise we can conclude that the optimum trajectory consists of intervals of maximum thrusting separated by intervals of zero thrust (coast arcs). The optimum pointing angle during thrusting can be determined.

6.3.2 Parameterized Optimization in Three Dimensions

Consider now the problem in three dimensions of optimizing the transfer from heliocentric injection to rendezvous with a remote planetary body, a planet, asteroid, or comet. Assume that the rendezvous is to occur at a given time, for example at aphelion of a comet when such a body is inactive and not surrounded by dust and debris which would be potentially harmful to a spacecraft. The problem delineated above is extended by (a) defining the 6 terminal conditions at time t_f to be the components of position and velocity of the target body, and (b) using three control variables to specify the thrust vector. A 7-component state vector x is defined in terms of the orbital elements $(a, e, \theta^*, \omega, i, \Omega)$ and the spacecraft mass m . The thruster acceleration vector is denoted as (f_t, f_n, f_z) in the "Tangential-Normal" (tnz) axes employed in section 3.2.5. The state equations of that section are repeated here with minor changes due to the use of θ^* instead of θ .

$$\begin{aligned} \frac{da}{dt} &= 2a^2 v f_t / \mu \\ \frac{de}{dt} &= \frac{1}{v} \left[2(e + \cos \theta) f_t - \frac{r}{a} \sin \theta f_n \right] \\ \frac{d\Omega}{dt} &= \frac{r \sin \theta^* f_z}{h \sin i} \\ \frac{d\omega}{dt} &= -\cos i \frac{d\Omega}{dt} \\ &\quad + [2 \sin \theta f_t + (2e + r \cos \theta/a) f_n] / ev \\ \frac{di}{dt} &= r \cos \theta^* f_z / h \\ \frac{d\theta^*}{dt} &= h/r^2 - \cos i \frac{d\Omega}{dt} \\ \frac{dm}{dt} &= -f_p u_1 m / u_i \end{aligned} \tag{6.15}$$

where

$$\theta = \theta^* - \omega \tag{6.16}$$

velocity v is given by eqn 3.26, and r by eqn 1.24.

The thruster accelerations with respect to tnz axes are quoted in terms of the polar angles γ (in-plane) and β (out-of-plane)

$$\begin{aligned}
 f_t &= f'u_1 \cos \beta \cos \gamma \\
 f_n &= f'u_1 \cos \beta \sin \gamma \\
 f_z &= f'u_1 \sin \beta
 \end{aligned} \tag{6.17}$$

where the dimensionless magnitude (nominally unity) satisfies

$$0 \leq u_1 \leq u_{max} \tag{6.18}$$

The acceleration factor f' is related to the electrical power by eqn 6.2. Thus if the available power from solar cells decays inversely as the power k from the Sun, the nominal acceleration is

$$f' = \frac{2\eta P_0}{u_i m R^k} \tag{6.19}$$

where R is distance from the Sun in AU, and u_i is exhaust velocity (gI_{sp}). Representative values for Solar Electric Propulsion are as follows (Ref.[42]) if f' is in m/sec².

P_0	34546 watts
η	0.61
u_i	37602 m/s
m	4388 kg (initially)
k	1.6

Detailed optimization studies for the rendezvous problem have shown that optimum thrusting is achieved with one coast arc as follows, t_0 being the start time.

u_1	Start time	End time
u_{max}	t_0	t_1
0	t_1	t_2
u_{max}	t_2	t_f

These studies employed iteratively a Quadratic Programming Algorithm (Ref. [17]) but other nonlinear programming routines for constrained optimization have also been applied (Ref. [19]).

As a simplified illustration of the optimization process the problem has been formulated parametrically by approximating the pointing angles as changing linearly with time t . Thus

$$\gamma = \gamma_0 + \gamma_1 t \tag{6.20}$$

$$\beta = \beta_0 + \beta_1 t \tag{6.21}$$

Given the assumption of one coast interval as above, we have consequently 8 parameters

$$t_0, t_1, t_2, u_{max}, \gamma_0, \gamma_1, \beta_0, \beta_1 \tag{6.22}$$

that can be adjusted to satisfy the 6 terminal conditions, viz. the position and velocity vector of the target body at the pre-specified time t_f . Notice that the total duration of the mission is adjustable via the start time t_0 . There are two degrees of freedom and therefore we can select 6 from the 8 parameters

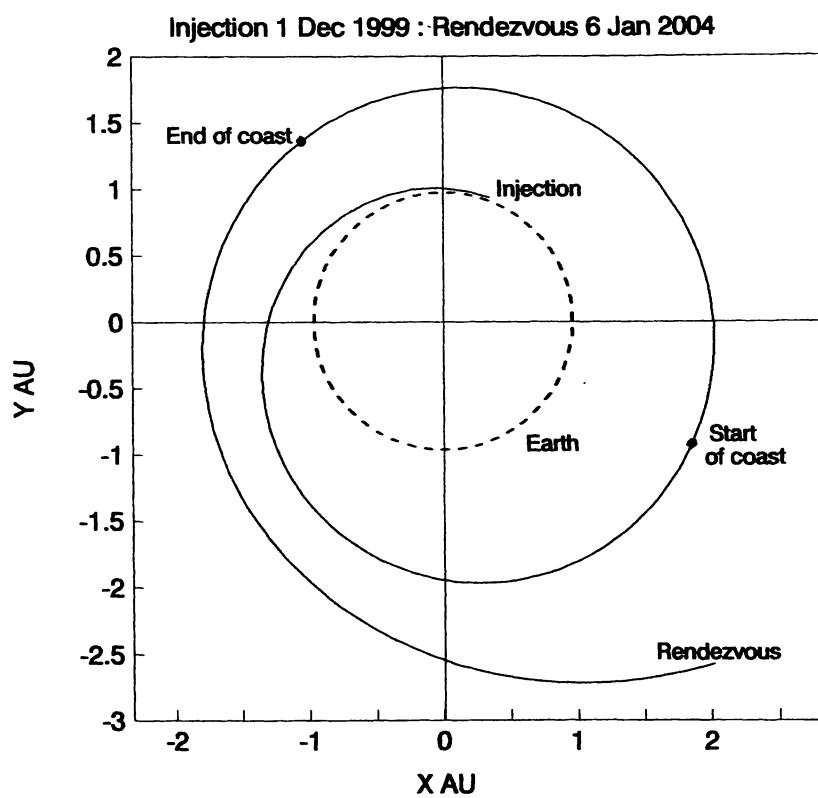


Fig. 6.2. Solar Electric Propulsion for Comet Rendezvous

(6.22) to satisfy the terminal conditions, leaving two parameters to maximize the final mass of the spacecraft. By the way, the choice of the pointing angles γ and β with respect to tnz axes is intentional, because we would expect these angles to be fairly small; the energy increases most rapidly when the thrust vector is aligned with the velocity vector.

6.3.3 Numerical Example

Software (section 6.5) was written first to simulate the spacecraft trajectory from injection at time t_0 to rendezvous at time t_f , dependent on the 8 parameters at eqn 6.22. In fact it was parameterized in one other respect to correspond to applying the hyperbolic escape vector at an angle ξ to the ecliptic plane. Thus, if the velocity vector of the Earth around the Sun is approximated as

$$\mathbf{V} = V(\cos \phi, \sin \phi, 0) \quad (6.23)$$

and the escape velocity vector is

$$\mathbf{V}_\infty = V_\infty(\cos \xi \cos \phi, \cos \xi \sin \phi, \sin \xi) \quad (6.24)$$

then the injection velocity vector is $(\mathbf{V} + \mathbf{V}_\infty)$. The above angle ϕ (ecliptic longitude) is adjusted by varying the date of the launch. A zero value for injection at t_0 corresponds to Julian days 2451542.0 (29 Dec 1999). In this example rendezvous is supposed to occur at the aphelion of a comet having approximately the parameters of *Bus* at Julian days 2453011.0 (6 Jan 2004), when the position and velocity vectors are respectively (2.017, -2.574, 0.118) AU and (16.270, 4.782, -0.195) km/s. The escape velocity V_∞ was taken to be a representative 2.500 km/s corresponding approximately to an increment from a parking orbit of 3.5 km/s and the angle ξ was adjusted and fixed at 17.2 degrees.

The state equations of the orbital elements 6.15 were integrated with a fourth order Runge-Kutta process, although the final state has to be transformed into the corresponding vectors of position and velocity, ie. for comparison with the desired values. The osculating orbit of this example is very close to the ecliptic plane. Therefore, to avoid the singularity at zero inclination, axes rotated 90 degrees about the Y axis were employed inside the integration routine. If \mathbf{p} is the vector of 6 chosen parameters and \mathbf{y} is the output vector of position and velocity at time t_f , then for small changes of the parameters

$$\delta \mathbf{y} = \Phi(t_f, t_0) \delta \mathbf{p} \quad (6.25)$$

The state transition matrix Φ has to be generated by applying small numerical perturbations to 6 repeated integrations of the trajectory from time t_0 to t_f . Then, if $\delta \mathbf{y}$ is the required correction in the output vector, the parameter vector is improved according to

$$\delta \mathbf{p} = \Phi^{-1} \delta \mathbf{y} \quad (6.26)$$

In practice, unless the first estimate is very good, the iterative step of eqn 6.26 has to be restricted by a factor such as 0.2 or 0.5 to ensure convergence.

Since there are 6 parameters to be adjusted out of a possible 8 of eqn 6.22, there are several options. However, it was found that *only the following set of 6 parameters would permit reliable convergence.*

$$(u_{max}, t_0, t_1, t_2, \beta_0, \beta_1) \quad (6.27)$$

Attempts to include γ_0 or γ_1 (associated with in-plane pointing of the thrust vector) resulted in big iterative adjustments and divergence. Thus, the final equality constraints could be satisfied with a range of values of γ_0 and γ_1 , which leaves the latter parameters available for optimization, eg. by a simple trial-and-error process. The following results summarize the end results of such an approximation to constrained optimization. Depending on the initial values of the 6 parameters, about 20 iterations were needed to satisfy the terminal equality constraints.

Optimization parameter	Iterated parameter	Value
γ_0		0.200 radian
γ_1		0.100×10^{-8} radian/sec
	u_{max}	1.068
	t_0	-28 days
	t_1	493 days
	t_2	893 days
	β_0	0.645 radian
	β_1	-0.646×10^{-8} radian/sec

The trajectory corresponding to the above parameters is illustrated at Figure 6.2. The final mass was 3221.5 kg, compared to 3279 kg resulting from an accurate constrained optimization using sequential quadratic programming (Ref. [17]).

6.3.4 Controllability

It was observed above that, although the two in-plane angular pointing parameters γ_0 and γ_1 could be adjusted to optimize the final result, large changes and divergence occurred if, for example, throttling (u_{max}) was replaced by γ_0 as a parameter to be adjusted to satisfy the end conditions. This raises the question as to how we can test systematically which parameters are effective in satisfying the terminal conditions.

In forming the Φ matrix of eqn 6.25 assume that all inputs and outputs of the matrix have been scaled to be of the same order of magnitude, eg. (t_0, t_1, t_2) are measured in thousand days and γ_1 has a scaling factor 10^{-8} . Apply now a *singular value decomposition* to the matrix Φ . As shown in standard texts on linear algebra, this yields the following result.

$$\Phi = UDV^T \quad (6.28)$$

If generally Φ has m rows and n columns ($m \leq n$), then

1. the orthogonal mxm matrix U consists of the m eigenvectors of $\Phi\Phi^T$.
2. the orthogonal nxn matrix V consists of the n eigenvectors of $\Phi^T\Phi$.
3. the mxn matrix D has zero elements except for the m *singular values* σ_i on the diagonal.
4. the singular values σ_i are equal to the square roots of the real and non-negative eigenvalues of the mxm symmetric matrix $\Phi\Phi^T$.

The required computer algorithm is available in most libraries of numerical methods although, if we are interested only in the singular values, use an eigenvalue program to compute the eigenvalues of $\Phi\Phi^T$. Define the well behaved choice of parameters to be $(u_{max}, t_0, t_1, t_2, \beta_0, \beta_1)$ to be option 1; a troublesome choice was option 2 viz. $(gamma_0, t_0, t_1, t_2, \beta_0, \beta_1)$. The following singular values were computed.

option 1	258	20.3	7.24	3.13	0.69	0.27
option 2	241	20.2	5.39	2.40	0.50	0.010

The singular value decomposition is a method of testing how the parameter space is mapped into the space of the output vector. If one of the singular values is zero then the matrix Φ has rank 5 (instead of 6) and it will be impossible to adjust the parameter vector to satisfy the end conditions. When one or more of the singular values is very small (assuming sensible scaling) then inversion of the matrix Φ will be very sensitive to small errors, eg. due to linearization. The better behaviour of option 1 versus option 2 is confirmed above by observing that the smallest singular values are 0.27 and 0.010 respectively. Another way of stating this property is to assert that throttling the low-thrust rocket motor is a much more effective way of guiding the spacecraft to desired terminal conditions than varying the in-plane pointing angle. The *controllability* of the latter is said to be poor. The out-of-plane angles are similarly weak control variables but there is no alternative to their use in correcting errors out of the ecliptic plane.

6.4 Guidance on Low Thrust Trajectories

6.4.1 Introduction

Let us assume that a nominal (near optimum) low thrust trajectory has been computed before a mission. The thruster on-off times are stored onboard the spacecraft as are the pointing angles. Assume that the latter have been converted into nominal angles $\bar{\phi}$ and ψ_b with respect to the ecliptic plane, since such reference axes would be available (indirectly) as part of the attitude determination system of the spacecraft. Thus the pre-computed nominal thrust vector would be

$$\bar{u} = u_{nom}(\cos \psi_b \cos \phi_b, \cos \psi_b \sin \phi_b, \sin \psi_b) \quad (6.29)$$

where u_{nom} has replaced the u_{max} of the previous section. This could be stored at intervals with linear or perhaps quadratic interpolation provided to determine the angles at intermediate points.

The actual trajectory of the spacecraft would differ from the pre-computed orbit principally for the following reasons.

1. The injection conditions would be subject to dispersion arising mainly from off-nominal performance of the last rocket stage.
2. The thrust developed by the electric propulsion system would vary from the nominal level, due to the efficiency of the solar cells, etc.
3. The direction of the thrust vector would be in error due to uncertainties in the centre of mass and imperfections of the spacecraft attitude determination and control system.
4. The desired final position and velocity will be updated in the late phase of the mission as the ephemeris of the target (comet) is improved by means of optical sightings from the spacecraft.

In-flight corrections will therefore be necessary, ie. a guidance system. Assuming ground-based operations, three procedures seem feasible.

1. Periodic re-optimization using software for constrained optimization, ie. a more advanced version of the technique of section 6.3.
2. Feedback guidance by varying the thrust pointing angles and on-off times.
3. Feedback guidance by varying the throttle level of the low thrust motor and the pointing angles.

A study by Noton and Salehi (Ref. [42]) for the European Space Agency concluded that the more complicated option 1 above was not necessary and that option 3 was much more effective than option 2. The reason was that controllability was poor without the use of motor throttling, ie. *exactly as illustrated by the computations of the previous section*. This option is therefore developed and illustrated here.

6.4.2 Analysis

A 7-component state vector x has already been introduced, viz. the orbital elements $(a, e, \theta^*, \omega, i, \Omega)$ and the spacecraft mass m , the state equations being listed at eqn 6.15. We propose to superimpose on the nominal 3 control variables (obtained from onboard interpolated storage) small corrective terms based on periodic updating of the spacecraft orbit from ground-based radiometric data. No such correction is of course possible in the coast phase but let the two thrusting phases be subdivided into equal intervals h , not necessarily the same for the two phases. If the state equations are integrated from time t_k to time t_{k+1} , we denote that process by

$$x(k+1) = f(x(k), \bar{u}) \quad (6.30)$$

where \bar{u} indicates use of the nominal control vector, ie. throttle level u_{nom} , and pointing angles (ϕ_b, ψ_b) .

It is proposed to employ a corrective control vector which is constant during any one subinterval h ; it is denoted $u(k)$ with components corresponding to changes in the thrust level, and angles ϕ and ψ . For small departures from the nominal trajectory, and provided $u(k)$ is sufficiently small, eqn 6.30 can be approximated by

$$\delta x_k + 1 = F(k)\delta x(k) + E(k)u(k) \quad (6.31)$$

where the 7×7 matrix $F(k)$ and the 7×3 matrix $E(k)$ can be generated by perturbed numerical integrations. $F(k)$ is the state transition matrix over the interval t_k to t_{k+1} . If desired, equation 6.31 can be extended to include the possibility of small additive random perturbations. Thus

$$\delta x_k + 1 = F(k)\delta x(k) + E(k)u(k) + C(k)w(k) \quad (6.32)$$

where $w(k)$ is a vector of independent random variables with zero mean and unit variance. Correlated random disturbances can be handled by augmenting the state vector with additional states. For example

$$x_8(k+1) = b_8 x_8(k) + c_8 \sqrt{1 - b_8^2} w(k) \quad (6.33)$$

with $b_8 = \exp(-h/T_c)$ represents a stochastic process with standard deviation c_8 and auto-correlation time T_c .

6.4.3 The Linear Quadratic Control Problem

Given the linear state equations with control at eqn 6.31 where ($k = 0, 1 \dots N$), suppose that we wish to choose the control variables $u(k)$ to minimize a quadratic performance index of the form

$$\begin{aligned} V(\delta x(0), u) &= \delta x(0)^T Q(N) \delta x(N) \\ &+ \sum_{k=0}^{N-1} [\delta x(k)^T Q(k) \delta x(k) + u(k)^T R(k) u(k)] \end{aligned} \quad (6.34)$$

This is the linear quadratic control problem of optimal control theory. The matrices $R(k)$ and $Q(k)$ are symmetric and positive definite although, if the matrices $R(k)$ are not strictly positive definite, then large but finite control corrections will arise. The solution to this problem is given in standard texts on or including optimal control theory (Refs. [8], [40], [20]). Therefore the solution is quoted without proof below; it is incidentally mathematically *dual* to the solution for the Kalman Filter (section 4.4.4).

The following equations are iterated by stepping backwards from $k = N - 1$ to $k = 0$, initiated by setting $G(N) = Q(N)$.

$$K(k) = G(k+1)E(k) [R(k) + E(k)^T G(k+1)E(k)]^{-1} \quad (6.35)$$

$$P(k) = G(k+1) - K(k)E(k)^T G(k+1) \quad (6.36)$$

$$G(k) = F(k)^T P(k)F(k) + Q(k) \quad (6.37)$$

For any subinterval k the control vector is given by

$$u(k) = -K(k)^T F(k) \delta x(k) \quad (6.38)$$

i.e. the control corrections are calculated as a linear combinations of departures of the state variables from the pre-computed nominal (optimum) trajectory.

The state equations were extended at eqn 6.32 to include stochastic terms. If the additive random terms can be approximated as Gaussian, then a further result follows from optimal control theory (Ref. [40]), viz. that the optimal control in the presence of such stochastic inputs (with zero mean value) is the same as if the stochastic inputs were set to zero. Therefore the above solution remains valid in the presence of such stochastic inputs. The state vector must naturally be estimated from measurements but the so-called *Separation Theorem* (Ref. [40]) states that optimal control is realised by implementing control based on the above linear quadratic formulation, but employing estimates of the state from a sequential Kalman estimator.

6.4.4 Numerical Example of Feedback Guidance

The matrix $Q(N)$ of eqn 6.34 penalizes final errors in the states and was set equal to the unit matrix except

$$Q_{11}(N) = w_1^2 \quad (6.39)$$

to permit some separate adjustment on the weight associated with the first state (a in AU). Bear in mind that, with the angles in radians and the mass in tonnes, all the states are the same order of magnitude. The penalty on previous deviations from the nominal trajectory was set to zero, ie. $Q(k)$ equal to the zero matrix for $(k = 0, 1\dots N-1)$. This represents therefore *end-point guidance*. In order to limit changes in the control variables

$$R(k) = \text{diag}(w_u^2, w_a^2, w_a^2) \quad (6.40)$$

For any pre-computed low thrust trajectory the 7x3 matrix $K(k)$ of eqn 6.38 can be computed and stored for operational use, or in a simulation. It would not be necessary to re-compute it unless very large unexpected changes occurred during an actual mission. To form the departures $\delta x(k)$ from the nominal state, the *estimated state* would be employed, ie. from the orbit determination. Remember that either a sequential method must be used, because the orbit changes continuously due to thrusting, or the batch size must be no longer than the control subinterval.

The weights for this illustrative simulation were chosen by trial and error as ($w_1 = 0.01$, $w_u = 1$, $w_a = 1$) and there were 50 subintervals in each of the two thrust phases, roughly 10 day subintervals. The results of three type of simulation are presented in Tables 6.1, 6.2, and 6.3 respectively as follows.

1. Disturbances due to random variations in thrust level of standard deviation 2 per cent, with the auto-correlation time constant set to 100 days. The priming number randomizes the start of the random number generator.
2. Pointing of the thrust vector subject to constant bias errors as listed in ϕ (in-plane angle) and ψ (out-of-plane angle).
3. Typical initial dispersion at heliocentric injection is randomized with RSS position errors equal to 10,000 km and RSS velocity errors equal to 25 m/s. The priming number again corresponds to a different sequence of random numbers.

Table 6.1. Guidance subject to Random Thruster Variations

Random priming	Final Errors		Extremes of Control Variables		
	RSS km	RSS m/s	thrust per cent	ϕ deg	ψ deg
0	938	0.3	12.4	0.8	2.4
1	589	0.5	11.8	0.9	2.3
2	1225	0.6	10.5	1.1	2.3
3	1331	0.2	10.4	1.3	2.2

Table 6.2. Guidance subject to Pointing Bias Errors

Bias deg	Final Errors		Extremes of Control Variables		
	RSS km	RSS m/s	thrust per cent	ϕ deg	ψ deg
$\phi = +0.5$	2269	0.3	0.8	2.3	1.2
$\phi = -0.5$	1884	0.3	1.1	2.0	0.9
$\psi = +0.5$	7346	0.4	1.3	0.6	2.3
$\psi = -0.5$	3442	0.5	1.3	0.6	2.3

Table 6.3. Guidance subject to Random Initial Dispersion

Random priming	Final Errors		Extremes of Control Variables		
	RSS km	RSS m/s	thrust per cent	ϕ deg	ψ deg
0	57	0.0	0.2	0.2	0.8
1	11	0.0	0.2	0.1	0.2
2	16	0.0	0.3	0.3	0.5
3	16	0.0	0.1	0.2	0.5

Referring to Tables 6.1, 6.2, and 6.3, in all cases there are *no errors due to orbit determination*; an adequate representation of these effects requires a parallel simulation of the full orbit determination process. While this represents another source of errors, Noton and Salehi (Ref. [42]) have demonstrated that errors in the thrust level and pointing angles can be calibrated as part of the orbit state estimation. The larger errors, for example due to pointing bias terms (Table 6.2), are therefore not very relevant. A separate study of *final guidance* would allow for the inclusion of optical sightings of the comet in the final approach to the latter. Such extra measurements permit big improvements in the estimated position and velocity of the spacecraft relative to the comet.

6.5 Software

The simulations of this section were carried out by means of C++ programs grouped together in the files SEP.CPP (for low thrust trajectory optimization) and SEG.CPP (low thrust guidance). All the routines are in double precision and linking to the previously introduced subroutines in ORBIT2.CPP and MATRIX2.CPP is necessary. Routines in those files are not listed below.

The routines included in the file SEP.CPP are as follows.

Main : main program to illustrate approximate optimization of a low thrust trajectory for rendezvous with a planetary body.

low_thrust : given vector of adjustable parameters, the output vector of position and velocity is computed.

Earth_RV : position and velocity of Earth given Julian days.

Runge_Kutta4 : Numerical integration by the fourth order Runge Kutta process.

RHS : Generates right hand sides of the state equations in orbital elements.

The routines included in the file SEG.CPP are as follows.

Main : main program to illustrate guidance about a pre-computed low thrust trajectory generated by the program SEP above.

lowthrust_guid : depending on an option number, this routine is used to control

1. storage of the nominal pointing angles.
2. computation and storage of the matrices of partial derivatives.
3. simulation of feedback guidance.

LQG : computation and storage of the control matrices.

random_normal : random numbers approximating a normal distribution.

Earth_RV : see above in SEP.CPP .

step_states : steps states through one subinterval using numerical integration if thrusting, or orbital routines if coasting when the subinterval is large.

Runge_Kutta4 : see above in SEP.CPP .

RHS : as above in SEP.CPP except the equations are augmented to include interpolating pointing angles from stored arrays and calculation of feed-back corrective control terms.

KX_from_store : control gain matrix and nominal state vector recovered from global arrays.

FE_from_store : matrices of partial derivatives recovered from global arrays.
store_point : store pointing angles in global arrays.

Transform : rotational transformation from ecliptic to tnz axes; see section 3.10.

CHAPTER 7

ATMOSPHERIC ENTRY

7.1 Introduction

Some space missions culminate in the entry of the atmosphere of either the Earth or a planet. Examples are as follows.

1. Return to Earth of the NASA Apollo lunar missions.
2. Return to Earth of all Shuttle missions.
3. Entry into the atmosphere of Titan (a satellite of Saturn) at the conclusion of the Cassini mission.
4. Return of samples to Earth after a visit to a comet (the original plan for ESA's Rosetta mission).

Re-entry would also be a critical phase of a possible future spaceplane for inter-continental travel involving a partial satellite orbit.

While the underlying dynamics is the same for all such missions, the influential parameters vary so much that the practical aspects of mission design differ greatly. For example, after de-orbiting from a low circular Earth orbit, the Shuttle will enter the atmosphere at about 7.8 km/s compared to about 15 km/s for a spacecraft returning from a deep space mission out to several astronomical units. Such differences have a disproportionate effect on the resulting atmospheric deceleration and heating. Decelerations on manned missions must be limited to about 4g whereas structures of unmanned capsules can be designed to tolerate perhaps as high as 60g. Lower levels of heating can be handled by special surfaces and insulating layers, but more intensive heating requires ablative surfaces which dissipate heat as they pass from solid to the vapour state.

This chapter reviews the equations of motion, expressions for aerodynamic lift and drag, and summarizes approximate heating formulae. The concept of *entry corridors* is then examined. With the help of computer software, ballistic unguided non-lifting entry is then compared to guided entry which exploits available lift, whether the latter arises from a winged vehicle or merely the low lift-to-drag ratio of an offset cone shaped capsule. Attitude control of lifting vehicles is not discussed.

7.2 Flight Equations

7.2.1 Equations of Motion

The equations of motion have already been derived in chapter 2 but they are re-stated here as a matter of convenience. The six dependent variables are as illustrated in Figure 2.1, viz.

- τ longitude
- δ latitude
- r radius from the centre of the Earth or planet
- v velocity
- χ azimuth angle (from N) of velocity vector
- γ elevation angle (above local horizontal)
of velocity vector

If w and a are the vectors in wind axes (Figure 2.2) of the gravitational and aerodynamic forces respectively, then the six first order differential equations are (for vehicle mass M)

$$\dot{\tau} = (v/r) \cos \gamma \sin \chi (\cos \delta)^{-1} \quad (7.1)$$

$$\dot{\delta} = (v/r) \cos \gamma \cos \chi \quad (7.2)$$

$$\dot{r} = v \sin \gamma \quad (7.3)$$

$$\dot{v} = \omega_e^2 r \cos \delta (-\cos \gamma \cos \chi \sin \delta + \sin \gamma \cos \delta) + (w_1 + a_1)/M \quad (7.4)$$

$$\begin{aligned} \dot{\chi} &= 2\omega_e (-\tan \gamma \cos \chi \cos \delta + \sin \delta) + \left(\frac{\omega_e^2 r}{v \cos \gamma} \right) \sin \chi \cos \delta \sin \delta \\ &+ (v/r) \cos \gamma \sin \chi \tan \delta + (w_2 + a_2)/(Mv \cos \gamma) \end{aligned} \quad (7.5)$$

$$\begin{aligned} \dot{\gamma} &= 2\omega_e \sin \chi \cos \delta + (\omega_e^2 r/v) \cos \delta (\sin \gamma \cos \chi \sin \delta + \cos \gamma \cos \delta) \\ &+ (v/r) \cos \gamma + (w_3 + a_3)/(Mv) \end{aligned} \quad (7.6)$$

The components of the gravitational and aerodynamic forces in wind axes are now required. If $g(r, \delta)$ is the acceleration due to gravity as a function of Earth radius and latitude, the vector of gravitational forces is

$$\mathbf{w} = (-g(r, \delta) \sin \gamma, 0, -g(r, \delta) \cos \gamma) M \quad (7.7)$$

The gravitational acceleration in expansion terms up to second order is (section 5.2, Ref. [57])

$$g(r, \delta) = \mu/r^2 [1 + 3J_2(R_0/r)^2(1 - 3\sin^2 \delta)/2] \quad (7.8)$$

where R_0 is the mean equatorial radius. See Table 2.2 for numerical values. The Earth radius as a function of latitude is given by (section 4.3, Ref. [57])

$$R(\delta) = R_0(1 - f \sin^2 \delta) \quad (7.9)$$

and therefore altitude is

$$h = r - R \quad (7.10)$$

7.2.2 Aerodynamic Lift and Drag

The lift vector is perpendicular to the velocity vector as shown in Figure 7.1 and the longitudinal X axis of the vehicle is at angle α (the angle of attack or incidence) to the velocity vector. However Figure 7.1 is simplified because all vectors are drawn as if they are in the vertical plane. In that case the lift vector in the wind axes of Figure 2.2 is $(0, 0, L(\alpha))$. Let us generalise this situation to allow for the vehicle being rolled angle β about the velocity vector, when the lift vector becomes $L(\alpha)(0, \sin \beta, \cos \beta)$. The total aerodynamic force a in wind axes is then

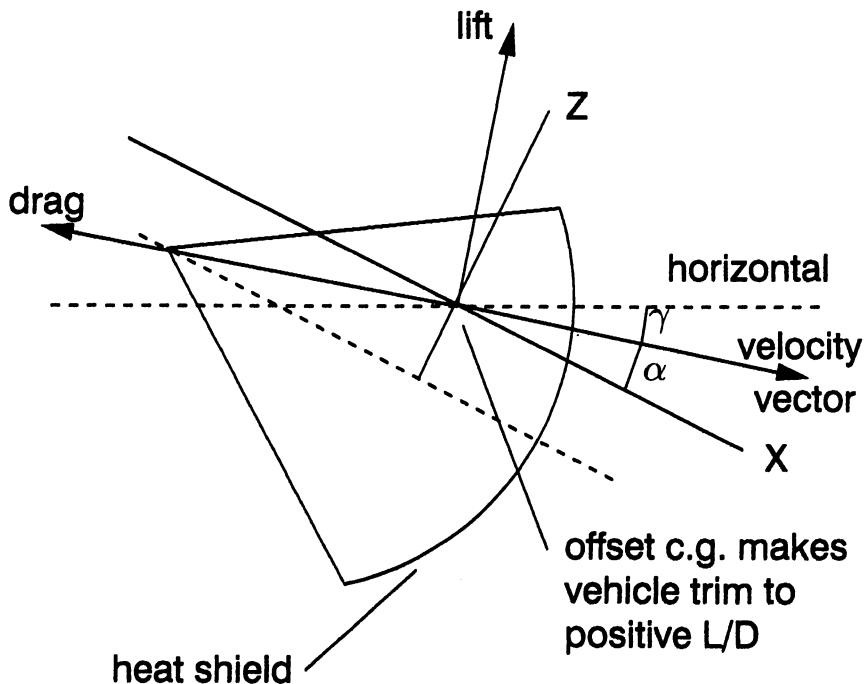


Fig. 7.1. Entry Capsule develops Lift

$$\mathbf{a} = D(\alpha) \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} + L(\alpha) \begin{bmatrix} 0 \\ \sin \beta \\ \cos \beta \end{bmatrix} \quad (7.11)$$

where the drag and lift coefficients (for a characteristic area S) arise as

$$D(\alpha) = \rho(h)v^2 C_D(\alpha)S/2 \quad (7.12)$$

$$L(\alpha) = \rho(h)v^2 C_L(\alpha)S/2 \quad (7.13)$$

The drag and lift coefficients can be approximated (independent of Mach number) by calculations such as those given by Regan (section 9, Ref.[45]). These were quoted in section 2.4.4 for a cone but they are required here for the typical shape of a re-entry vehicle illustrated in Figure 7.1, the so-called Apollo shape with a blunt nose. By reworking Regan's formulae for this shape the following results are obtained, where the characteristic area S is the greatest cross sectional area of that shape, and the semi-angle of the pointed cone of Figure 7.1 is equal to $(\pi/2 - \theta_c)$.

$$\begin{aligned} C_N &= (1/2) \cos^2 \theta_c \sin 2\alpha \\ C_A &= \cos^2 \alpha (1 + \sin^2 \theta_c) + (1/2) \sin^2 \alpha \cos^2 \theta_c \end{aligned} \quad (7.14)$$

and

$$\begin{aligned} C_D &= + \cos \alpha C_A + \sin \alpha C_N \\ C_L &= - \sin \alpha C_A + \cos \alpha C_N \end{aligned} \quad (7.15)$$

As an example, θ_c is equal to approximately 60 degrees for the heat shield of the Apollo re-entry capsule, leading to a calculated drag coefficient at zero angle of attack of 1.75. The value from experimental model tests (Ref. [14]) was 1.5 for the corresponding reference area. Notice that from inspection of equations 7.4, 7.12, and 7.15 that the *dominant aerodynamic deceleration* is proportional to the parameter

$$B_c = \frac{M}{C_D S} \quad (7.16)$$

where C_D is the drag coefficient, taken in this formula to be at zero angle of incidence. The parameter B_c is known as the ballistic coefficient, having the dimensions kg per sq.m. It is very roughly proportional to vehicle size, and is a very influential parameter in characterizing aerodynamic trajectories. The density of the atmosphere ρ appears in equations above. In the case of the Earth, reasonably accurate models are available as specified at the end of section 2.4.4. The atmospheres of planets are obviously not so well known although data improves with each space mission involving entry into their atmospheres. If the atmospheres of Mars, Venus, and Titan are represented *very approximately* by the formula

$$\rho = \rho_0 \exp\left(\frac{h_2 - h}{h_1}\right) \quad (7.17)$$

then representative values of the parameters are as listed in Table 7.1.

7.2.3 Heating

The total energy dissipated per unit time is equal to drag force multiplied by air velocity but almost all of this passes into the atmosphere. Heating of the surface of the vehicle occurs mainly due to convection; radiative heating becomes significant only above about 12 km/s. Computation of the convective

Table 7.1. Indicative Parameters of Planetary Atmospheres

Planet	ρ_0 kg/m ³	h_2 km	h_1 km
Mars	0.0019	20	9.8
Venus	10.5	29	8.5
Titan ¹	5.1	0	76

¹ Based merely on surface pressure 1.6 times that of the Earth and an atmosphere 10 times thicker.

heating over the whole surface of a vehicle is complicated but an approximate formula can be employed for the critical estimation of the heat flux over the hottest part of the vehicle surface, which is at the stagnation point near the nose. Assuming laminar flow this heat rate in watts per sq. m. for the atmosphere (of the Earth) is approximately

$$Q = 1.45 \times 10^{-4} \sqrt{\rho/R_n} v^3 \quad (7.18)$$

where v is air velocity and R_n is the radius of the nose. The reference would be an internal report but a very similar expression has been given by Chapman (Ref. [10]) and used in NASA simulations (Ref. [61]); see also Zoby (Ref. [62]).

In the absence of ablative material, almost all of the heat of eqn 7.18 is lost by radiation from the surface because the amount conducted inwards is relatively small. It follows from the law of radiation that

$$Q = \epsilon\sigma(T_s^4 - T_\infty^4) \quad (7.19)$$

where T_s and T_∞ are the surface and ambient temperatures, ϵ is the emissivity, and σ is the Stefan-Bolzman constant (5.67×10^{-8} watts/m² per deg K⁴). Current technology of surface materials requires that the surface temperature is limited to about 2000 deg K. Therefore, assuming emissivity equal to 0.8, use of the above yields the conclusion that the limit of a non-ablative system is about 73 watts per sq. cm. While winged vehicles such as the Shuttle can rely on an insulating system, hyperbolic entries into the atmospheres of the Earth or planets give rise to stagnation heating rates far in excess of this figure, and consequently ablative heat shields become necessary. Low density ablatives are competitive in the range 40-500 watts per sq. cm. but high performance ablatives are required above this range (Ref. [13]).

7.3 Entry Corridors

7.3.1 Parameters

Figure 7.2 illustrates schematically the approach to the Earth after returning from a planetary or lunar mission. The entry velocity of 14.4 km/s is

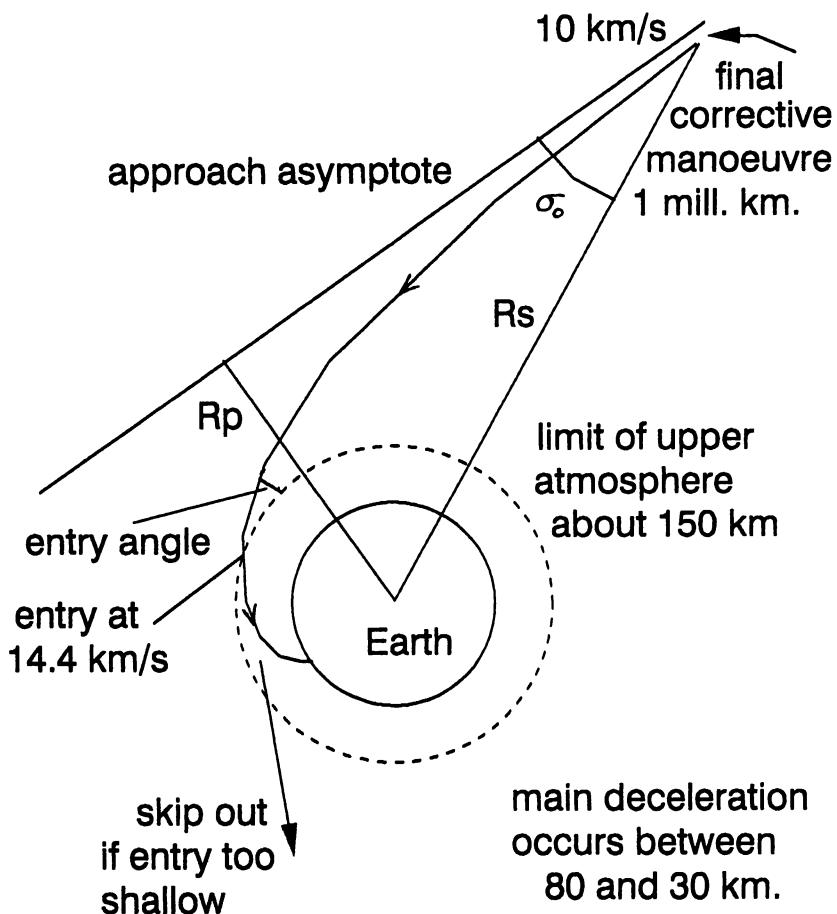


Fig. 7.2. Approach and Entry to the Atmosphere

representative for a deep space mission but would be lower for a return trip from the Moon, eg. 11 km/s. Outside the sphere of influence of the Earth the approach velocity vector is adjusted by a final midcourse manoeuvre to approach the Earth with an offset R_p , although the latter will be subject to errors in the orbit determination and in executing the corrective manoeuvre. If the approach is approximated as if from infinity (Fig. 7.2) then the entry conditions can be calculated from the energy equation 1.23 and the angular momentum.

$$V_\infty^2/2 = V_p^2/2 - \mu/R_{min} \quad (7.20)$$

$$h = R_p V_\infty = R_{min} V_p \quad (7.21)$$

leading to

$$R_p/R_{min} = \frac{\mu}{R_p V_\infty^2} + \sqrt{1 + \left(\frac{\mu}{R_p V_\infty^2}\right)^2} \quad (7.22)$$

The entry angle at radius r is obtained from

$$\begin{aligned} v^2 &= \dot{r}^2 + (h/r)^2 \\ &= V_\infty^2 + 2\mu/r \\ \dot{r} &= \tan \gamma (h/r) \\ h &= R_p V_\infty \end{aligned} \quad (7.23)$$

These equations yield

$$\cos \gamma = (R_p/r) \frac{1}{\sqrt{1 + 2\mu/r V_\infty^2}} \quad (7.24)$$

The entry angle is a critical determinant of the subsequent accelerations and heating imposed on the vehicle. In order to illustrate such conditions, calculations are summarized below for two sets of parameters corresponding to

1. An unmanned small space capsule returning from a deep space mission, eg. the original form of the *Rosetta* programme to bring samples from a comet back to Earth.
2. A manned vehicle returning from the Moon, corresponding approximately to the re-entry module of the NASA *Apollo* vehicle used famously for manned landing on the Moon.

Both vehicles have essentially the same "Apollo" shape as illustrated in Fig. 7.1. The angle between the vehicle X axis and the tangent at the extremity of the blunt nose is 60 degrees. Other parameters are summarized in Table 7.2.

Table 7.2. Parameters of Two Entry Capsules

Parameter	Small unmanned capsule	Manned Apollo re-entry module
ballistic coef.	72.8 kg/m ²	280 kg/m ²
max. total accn.	20g	4g
L/D	0.30	0.30
max. heating	500 watts/cm ²	60 watts/cm ²
nose radius	2.0 m	3.9 m

7.3.2 Condition for Level Flight

Note the equation 7.6 for vertical acceleration when the small terms associated with Earth rotation are ignored and flight is restricted to the vertical plane.

$$v\dot{\gamma} = (v^2/r) \cos \gamma - (\mu/r^2) \cos \gamma + (1/2)\nu\rho v^2/B_c \quad (7.25)$$

having substituted for the ballistic coefficient B_c (eqn 7.16) and introduced ν as the lift-to-drag ratio (L/D). The drag coefficient has also been approximated as that at zero incidence.

By setting γ and the left hand side of eqn 7.25 to zero, the condition for level flight can be obtained in the following form.

$$v = v_c \left[1 + \frac{\nu r \rho(h)}{2B_c} \right]^{-1/2} \quad (7.26)$$

where the velocity in circular orbit is

$$v_c = \sqrt{\mu/r} \quad (7.27)$$

Note that the density ρ is calculated from a table as a function of altitude h . If L/D is negative (lift down), the vehicle will ascend if the velocity exceeds that given by eqn 7.26 and vice versa. If L/D is positive (lift up), the vehicle will descend if that velocity is exceeded and vice versa. The two curves are shown in Figures 7.3 and 7.4 for the two illustrative vehicles respectively. Both plots reveal the same characteristics as follows.

1. In the upper atmosphere above about 75 km, even L/D equal to -0.3 (down lift) does not permit level flight. The entry angle must be sufficiently down to avoid skipping out.
2. Between altitudes 70 and 30 km when the velocity is likely to be in the range 10 to 2 km/s, the negative lift (L/D=-0.3) will result in a fairly steep descent but the vehicle will ascend if the lift is positive.
3. At lower altitudes below about 30 km the maximum positive lift (L/D=+0.3) is required for a gradual descent only when the air velocity is well below 1 km/s, ie. into the region of subsonic speeds.

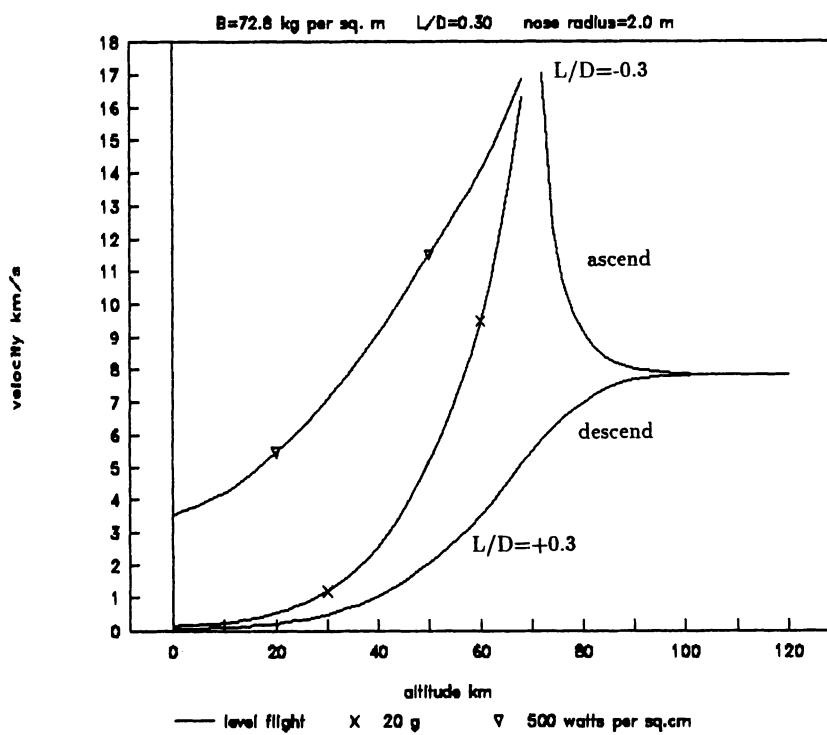


Fig. 7.3. Entry Corridors for a Small Unmanned Capsule

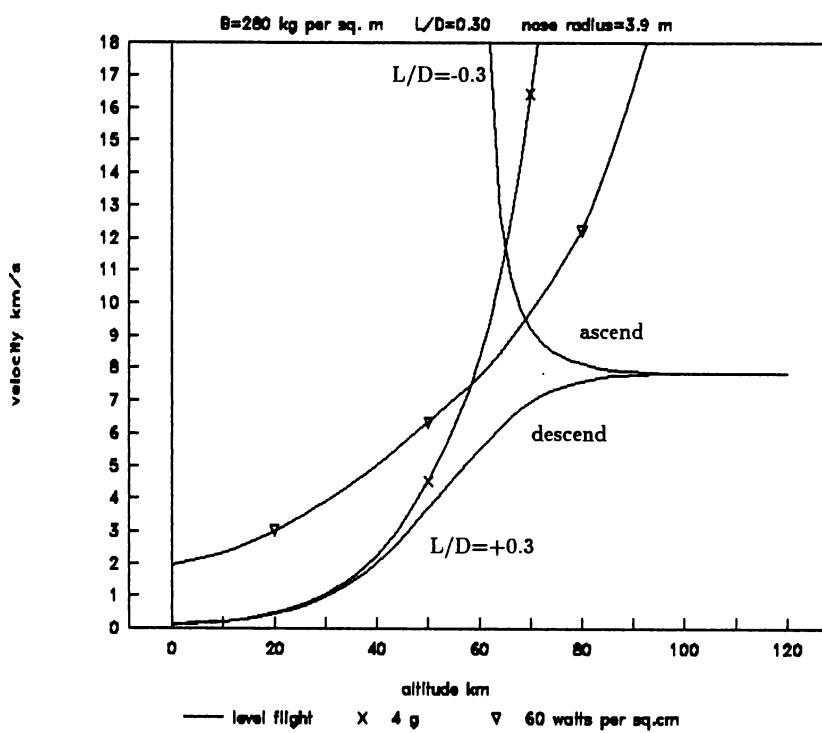


Fig. 7.4. Entry Corridors for Manned Re-entry Vehicle

7.3.3 The Limit of Total Acceleration

The acceleration experienced by the structure and crew (if any) is that due to aerodynamic forces. It is

$$\begin{aligned} a &= (1/2)\sqrt{C_L^2(\alpha) + C_D^2(\alpha)\rho V^2 S/M} \\ &= (1/2)\sqrt{1 + \nu^2} \rho v^2 C_D(\alpha) S/M \end{aligned} \quad (7.28)$$

ν being the lift-to-drag ratio. Approximate $C_D(\alpha)$ by C_D for small α and substitute for the ballistic coefficient (eqn 7.16), yielding

$$a = (1/2)\sqrt{1 + \nu^2} \rho v^2 / B_c \quad (7.29)$$

Conversely, if the altitude (hence density) and tolerable acceleration a_{max} is specified, the maximum velocity is given by

$$v_{max} = \left[\frac{2B_c a_{max}}{\rho(h)\sqrt{1 + \nu^2}} \right]^{1/2} \quad (7.30)$$

Figures 7.3 and 7.4 show the limiting curve respectively for 20g (representative for an unmanned capsule) and 4g (upper limit for a manned vehicle). At the higher velocities above about 2 km/s the altitude of the trajectory must not be too low, ie. into the denser atmosphere. The effect is to restrict the permissible combination of height and speed to the region between the three curves of maximum acceleration and level flight (not too steep ascent or descent) for $L/D=\pm 0.3$. The result is qualitatively the same for the two cases but the permissible region is much smaller for the manned vehicle, ie. *choice of trajectory and entry conditions will be more critical for manned entry.*

7.3.4 The Limit of Rate of Heating

The approximation for the rate of heating at the hottest point at the nose has already been quoted at eqn 7.18. If the maximum rate of heating Q_{max} is expressed in watts per sq. cm. and the altitude is given, then the maximum velocity (m/s) is obtained from

$$v^3 = \frac{Q_{max}}{1.45 \times 10^{-8} \sqrt{\rho(h)/R_n}} \quad (7.31)$$

R_n being the radius of the nose of the nose in metres and the density is in kg/m^3 .

These limiting curves are shown in Figures 7.3 and 7.4 for the two cases. As listed at Table 7.2 a high rate of 500 watts per sq. cm. was assumed for the unmanned capsule, corresponding to an ablative heat shield, whereas the lower figure of 60 watts per sq. cm. has been employed for the Apollo-type vehicle. It is clear that although the potential heating constraint is not relevant in this example for the small unmanned capsule, in the case of the manned vehicle below 50 km, the assumed much lower permissible rate of heating is a potential limitation.

7.4 Ballistic Entry with No Lift

As an example, numerical details are now presented for a representative ballistic entry, ie. unguided with no lift. It is the small unmanned capsule with the shape of Figure 7.1 but the c.g. is not offset and the lift-to-drag ratio is zero. Apart from the parameters already quoted in Table 7.2 above, the following were also employed in the computer simulation using the software listed in section 7.6.

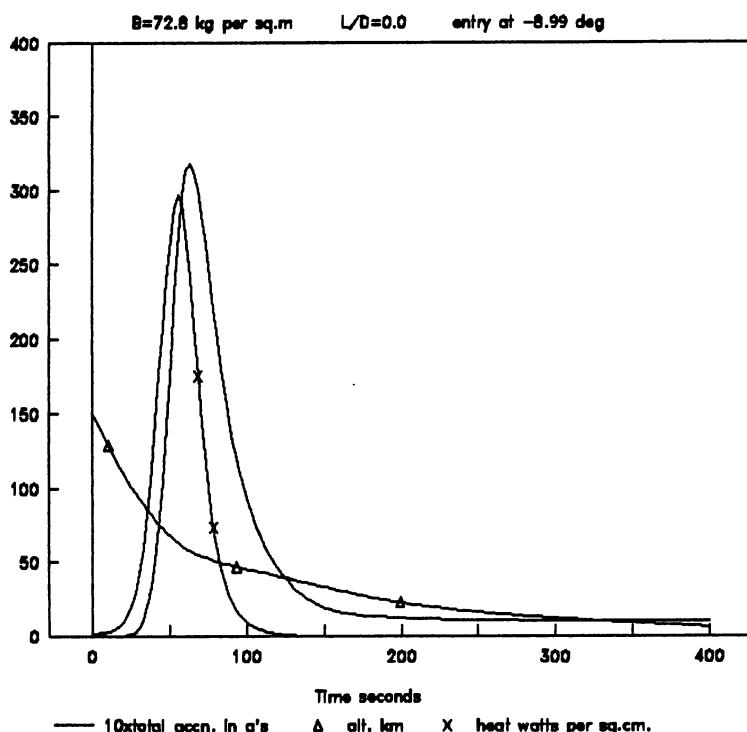


Fig. 7.5. Unguided Entry with no lift

mass	400 kg
cross-section radius	1.0 m
θ_c	60 deg.
$C_D(\alpha = 0)$	1.75
entry air velocity	14.41 km/s

The entry was in the equatorial plane which is not always achievable. It was an easterly entry, which would almost always be chosen because the air velocity is reduced by that of the Earth's rotation at the sidereal rate (0.47 km/s). For an entry of -8.99 degrees (at 150 km), Figure 7.5 shows plots of

altitude, total acceleration, and heat rate in watts per sq. cm. The scale for acceleration is in units of 10g, eg. the maximum is 31.8g. Note how most of the deceleration and heating takes place in the altitude range 80 to 30 km.

The maximum values of acceleration and heating are very sensitive to the entry angle as shown in Figure 7.6. The plot of acceleration is in m/s^2 , ie. divide by 9.81 for the figure in g's. For entries above (more positive) than -8.1 degrees the vehicle skips out; it escapes the atmosphere and departs on a modified hyperbolic orbit. Thus we can deduce from Figure 7.6 that if, for example, the acceleration is to be limited to 32 g and skip out is to be avoided, the entry angle must be restricted to the range -8.1 to -9.0 degrees. The corresponding approach offsets R_p are 9621 and 9600 km.

The final approach will have been corrected by a manoeuvre at typically 1 million kilometres. At this range when the velocity is 10 km/s, an error in the lateral velocity increment of 1 m/s introduces an error of 100 km in the offset R_p , or from eqn 7.24 a change in the entry angle of 3.8 degrees. It follows that, if the angular limits of the 32g entry corridor are to be satisfied (± 0.45 deg.), then the last lateral manoeuvre at 1 million km must be to an accuracy of ± 0.12 m/s. For example, if this manoeuvre can be executed to an accuracy of say no better than 2 per cent, then it must not be greater than 6.0 m/s. This may be sufficiently demanding that an earlier correction might be necessary, in order that the last manoeuvre is small enough.

These illustrative calculations underline the shortcomings of using an unguided ballistic entry but, apart from limiting the acceleration and heating loads, it would usually be necessary to terminate the entry (with a parachute) in a well defined area over land or water in order to permit quick rescue of the crew (if any) and recovery of the capsule. Using the above computed example of a ballistic entry, if the initial entry angle is restricted only to the ± 0.45 degree then the corresponding dispersion in the down-range distance is 1177 km. Given also some cross-range dispersion, the rescue and recovery operation would have to cover an area so enormous as to be unacceptable. Furthermore *weather might be bad in parts of such a big area* and there is no capability for diverting the touch-down with a ballistic entry. Ground tracking and orbit determination would continue up to entry but, during the subsequent severe atmospheric deceleration, communication with a transponder in the vehicle is not possible due to ionization. Thus data would be available to determine the actual entry angle and consequently estimate the expected point of touch-down but, since the duration of entry (in this example) is approximately 10 minutes, *there would be very little warning of where, in a large area, rescue should be concentrated*.

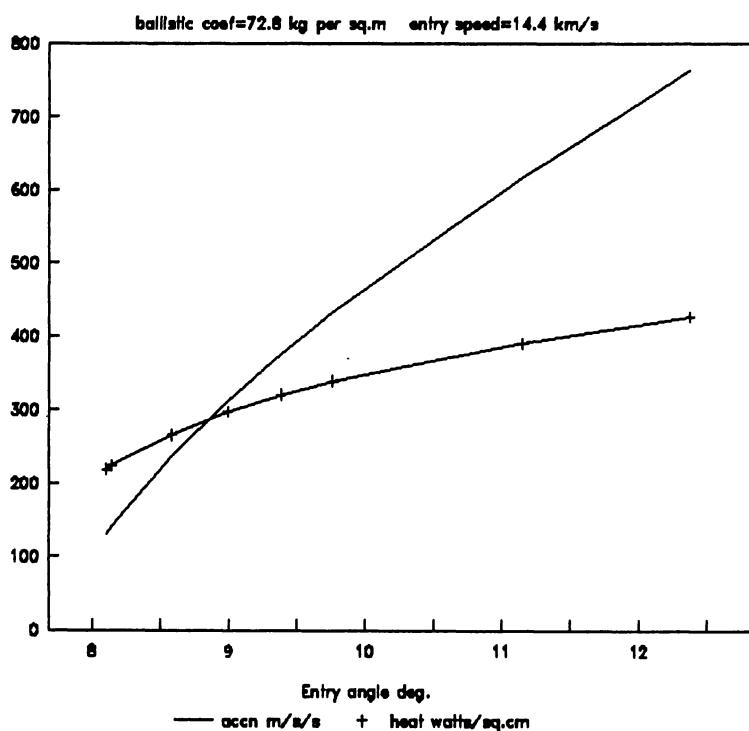


Fig. 7.6. Peak Acceleration and Heating for Ballistic Entry

7.5 Guided Entry using Lift

7.5.1 Lift and the Guidance System

Unlike a winged vehicle such as the Shuttle Orbiter, the type of entry capsule illustrated in Figure 7.1 develops lift only if the c.g. is offset from the longitudinal axis of symmetry of the vehicle. Referring to that diagram, let the X' axis be the longitudinal axis of the body, the broken line parallel to the X axis. Let the Z' axis be such that the origin is the aerodynamic centre of pressure. Assume that the centre of gravity is at (\bar{x}, \bar{y}) . The condition for equilibrium is obtained by setting to zero moments about the c.g. of the lift L and drag D vectors which operate through the centre of pressure. Thus

$$\bar{x}(L \cos \alpha + D \sin \alpha) - \bar{y}(L \sin \alpha - D \cos \alpha) = 0 \quad (7.32)$$

It follows that

$$\bar{y} = \bar{x} \frac{L/D + \tan \alpha}{L/D \tan \alpha - 1} \quad (7.33)$$

where, for static stability the c.g must be forward of the centre of pressure ($\bar{x} > 0$). If a given L/D is required, calculate the corresponding α from equations 7.14 and 7.15, hence the necessary offset \bar{y} from the longitudinal axis of the vehicle. The vehicle will then "trim" naturally to this orientation in flight, ie. a stable constant angle of incidence.

The lift vector can be directed by rolling the vehicle and indeed this is the only way that control of the trajectory can be achieved in both the vertical and horizontal plane. From eqn 7.11 the lift-to-drag ratio along the Z and Y wind axes (Figure 2.2) is $L/D \cos \beta$ and $L/D \sin \beta$ respectively, where β is the roll angle about the velocity vector. Since the descent is always relatively shallow, it follows that these are also approximately the vertical and horizontal components. The common method of realising guidance in the vertical and horizontal planes is as follows:

1. apart from pointing the lift vector up or down, the *mean lift in the vertical plane* can be reduced by rolling the vehicle symmetrically from side to side.
2. the mean lift in the horizontal plane can be adjusted by introducing non-symmetry (a bias) in the above rolling motion.

For example, suppose the roll angle is set (as a function of altitude h) according to

$$\beta(h) = \beta_0 [(1 - u_1/\beta_0) \text{sgn}(\omega_0 h) + u_2/\beta_0] \quad (7.34)$$

where sgn denotes a periodic square wave equal to ± 1 . The *mean vertical lift-to-drag ratio* is obtained by calculating the mean value of $\cos \beta$ over a multiple of the square wave as

$$(L/D)_v = L/D[\cos \beta_0 + \sin \beta_0 u_1] \quad (7.35)$$

where u_1 is small. Similarly, when u_2 is small, the *mean horizontal lift-to-drag ratio* is obtained by taking the mean value of $\sin \beta$ as

$$(L/D)_h = L/D[\cos \beta_0 u_2] \quad (7.36)$$

The value of the above β_0 can be used to set a nominal profile for L/D (positive negative or an intermediate value) and u_1 and u_2 can be employed for small corrections about the nominal setting.

In practical terms a system of guidance, navigation, and control is required onboard the vehicle. Control of attitude is required to orient the capsule before entry and to implement roll manoeuvres during the atmospheric phase. Navigation will determine the current position and velocity or equivalent state variables, and guidance will be employed to correct the trajectory. Attitude control of entry capsules is commonly by small thrusters, although larger vehicles have aerodynamic control surfaces for that phase. Navigation must be by an autonomous inertial system during actual atmospheric entry for reasons already given (ionization), although initialization would be available from the ground tracking network. Inertial navigation systems have been summarized in section 2.9; methods of guidance are now discussed.

7.5.2 The Nominal Trajectory

It is convenient to introduce the concept of a nominal entry trajectory despite the uncertainties in the actual entry angle, the aerodynamics, density in the upper atmosphere, and winds. However, as shown in Figure 7.4, constraints tend to dominate the choice of entry profile that must be followed. This is well illustrated in the description by Harbold and Graves (Ref. [22]) of the Apollo entry guidance, and by the same authors nine years later of the system for the Shuttle (Ref. [25]). More recent attention has been devoted to *optimizing* the choice of the entry trajectory, for example by a nonlinear programming algorithm (Refs. [12] and [27]).

By way of an introduction using simplified equations of motion, we will show how certain entry profiles can be estimated analytically. Thus, by ignoring small terms arising from the rotation of the Earth and restricting the analysis to a vertical plane, the equations of motion 7.3, 7.4, and 7.6 can be approximated by

$$\dot{h} = v \sin \gamma \quad (7.37)$$

$$\dot{v} = -\frac{\rho v^2}{2B_c} - g \sin \gamma \quad (7.38)$$

$$\dot{\gamma} = \frac{g \cos \gamma}{v} (v^2/v_c^2 - 1) + \frac{\rho v}{2B_c} (L/D) \quad (7.39)$$

where v_c is the circular velocity approximated as a constant equal to gR_0 , g being the acceleration due to gravity at the Earth's surface at radius R_0 . Drag deceleration is approximated by

$$d = (1/2)\rho v^2/B_c \quad (7.40)$$

Constant Rate of Descent. Suppose that during one phase of the entry the rate of descent is to be held constant equal to \dot{h}_0 , a negative value. From eqn 7.37

$$\gamma = \arcsin(\dot{h}_0/v) \quad (7.41)$$

Express eqn 7.39 as a solution for L/D but re-arrange it in the form

$$L/D = \frac{2B_c}{\rho v^2} (v\dot{\gamma} - g \sin^2 \gamma / \cos \gamma) + \frac{2B_c}{\rho v^2 \cos \gamma} (1 - \cos^2 \gamma v^2 / v_c^2) \quad (7.42)$$

Now $v \sin \gamma$ is constant (equal to \dot{h}_0), so by differentiation of it

$$v\dot{\gamma} = -\dot{v}\gamma \quad (7.43)$$

Therefore the first term in eqn 7.42 becomes

$$\begin{aligned} (v\dot{\gamma} - g \sin^2 \gamma / \cos \gamma) &= -\tan \gamma (\dot{v} + g \sin \gamma) \\ &= \tan \gamma \frac{\rho v^2}{2B_c} \end{aligned} \quad (7.44)$$

having substituted eqn 7.38. Equation 7.42 can consequently be written as

$$L/D = \tan \gamma + \frac{g}{d \cos \gamma} (1 - \cos^2 \gamma v^2 / v_c^2) \quad (7.45)$$

where the drag d has been introduced by means of eqn 7.40. Equation 7.45 reveals that, in order to descend at a constant rate, the lift-to-drag ratio should be adjusted as a function only of the drag deceleration and the velocity. This is *very convenient because both such variables would be available measurements*, ie. from accelerometers and integrated accelerometer outputs.

Constant Drag Deceleration. We now examine how the lift-to-drag ratio should be adjusted to maintain a constant level of drag deceleration d_0 . By differentiation of eqn 7.40, a constant,

$$r\dot{\rho}/\rho = -2\dot{v}/v \quad (7.46)$$

Assume that for this phase of the entry (eg. the upper atmosphere) the density can be approximated by

$$\rho = \rho_0 \exp(-h/h_s) \quad (7.47)$$

Then

$$\dot{\rho}/\rho = -\dot{h}/h_s \quad (7.48)$$

and consequently

$$\dot{h}/h_s = 2\dot{v}/v \quad (7.49)$$

Combine now eqns 7.37, 7.38 and 7.49 to yield

$$v^2 \sin \gamma / h_s = -\rho v^2 / B_c - 2g \sin \gamma \quad (7.50)$$

but re-arrange it as

$$(2g + v^2/h_s) \sin \gamma = -\rho v^2/B_c = -2d_0 \quad (7.51)$$

which must be constant by virtue of eqn 7.40 for constant drag deceleration. Differentiate the constant left hand side of eqn 7.51 with respect to time to give

$$v \dot{\gamma} = -\frac{2v^2 \dot{v} \sin \gamma}{\cos \gamma (2gh_s + v^2)} \quad (7.52)$$

When the constant drag deceleration is equated to d_0 in eqn 7.40, eqn 7.39 can be arranged as

$$L/D = [v \dot{\gamma} - g \cos \gamma (v^2/v_c^2 - 1)] / d_0 \quad (7.53)$$

Substitute eqn 7.52 into 7.53 to obtain

$$L/D = \frac{2 \tan \gamma v^2 (1 + g \sin \gamma / d_0)}{2gh_s + v^2} - g \cos \gamma (v^2/v_c^2 - 1) / d_0 \quad (7.54)$$

where γ can be eliminated by means of eqn 7.51 in the form

$$\gamma = \arcsin \left[\frac{-2d_0 h_s}{(2gh_s + v^2)} \right] \quad (7.55)$$

Equations 7.54 and 7.55 show how the *lift-to-drag ratio should be adjusted as a function only of velocity v*, which would be available as a measured variable from the inertial navigation system. There is one proviso arising from the differentiation of eqn 7.51. If subscript i denotes initial values we should have the initial state

$$\rho_i = \rho_0 \exp(-h_i/h_s) = \frac{2B_c d_0}{v_i^2} \quad (7.56)$$

or by taking logarithms

$$h_i = -h_s \log_e \left[\frac{2B_c d_0}{\rho_0 v_i^2} \right] \quad (7.57)$$

7.5.3 Feedback Guidance in the Upper Atmosphere

At the time of writing the best example of guidance during atmospheric entry is by means of the system developed in the 1960's by the MIT Instrumentation Laboratory for the returning Apollo lunar vehicles. Re-entry occurs now routinely for the Shuttle but this is after de-orbiting from circular orbit at about 7.8 km/s. Apollo entered at about 11 km/s which is almost twice as demanding in terms of deceleration, and our example is of a space probe entering the atmosphere at 14.4 km/s. The Apollo entry guidance system was fairly complicated with several phases (Ref. [22]) with two backup modes, but the essential underlying philosophy was as follows.

1. Dissipate most of the kinetic energy of the spacecraft in the upper atmosphere, almost allowing skip-out before restarting the descent. Adjustment of L/D (by rolling) in this phase is principally with respect to the required drag deceleration as a function of measured velocity.
2. Adjust L/D (by rolling) subsequently to divert, if necessary due to bad weather, and then achieve a specified point for splash-down within about 2 km.

Further research has naturally been devoted to entry guidance eg. optimization of the nominal entry and application of modern control theory to the feedback terms in the form of the LQ algorithm (Ref. [1]). The latter re-affirmed the importance of storing reference variables as a functions of parameters other than time, such as altitude or velocity. Nevertheless a re-design of the Apollo guidance system today would probably not be very different, certainly not in performance. A platform inertial navigation system would probably be replaced by a strapdown system with body mounted accelerometers and gyros. The onboard computer would be much more powerful, permitting more elaborate calculations in real time, but this is not a critical factor except perhaps with respect to redundancy and failure modes.

Returning to the example of the unmanned capsule coming back from a deep space mission, we simulate first guidance appropriate to the severe braking that occurs in the upper atmosphere. Note for example in Figure 7.5 how most of the deceleration occurs above about 40 km, and it is very similar with a low lift- to-drag ratio such as 0.3 which is now assumed. The vehicle parameters are as quoted in Table 7.2 and additionally as listed in section 7.4. The nominal lift-to-drag ratio is to be as calculated at eqn 7.54 above 50 km and according to eqn 7.45 below this altitude, ie. emphasis at the higher altitudes is placed on constant drag deceleration but at the lower altitudes on a constant descent rate. However, if we employed only this information for setting the lift-to-drag ratio, it would be "open-loop guidance". We will add feedback terms to correct for uncertainties in the atmosphere, the aerodynamics, but especially errors in adjusting the entry angle γ .

Given the importance of drag and altitude rate, because they would be available as measurements, and following practice in the Apollo and Shuttle systems, we choose the guidance law above 50 km to be

$$(L/D)_v = (L/D)_1 + c_1[a - a_{nom}(v)] + c_2[\dot{h} - \dot{h}_{nom}(v)] \quad (7.58)$$

and below 50 km as

$$(L/D)_v = (L/D)_2 + c_2[\dot{h} - \dot{h}_{nom}(v)] \quad (7.59)$$

with reference to the following notes.

1. L/D is limited to ± 0.3 and intermediate values of the vertical component $(L/D)_v$ are achieved by rolling from side to side as specified at eqn 7.35.

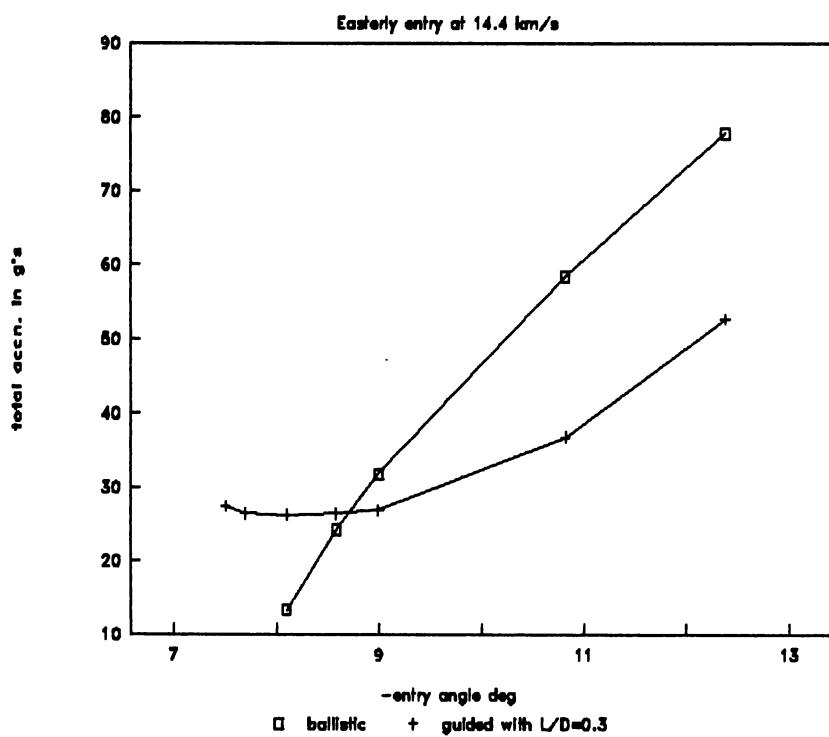


Fig. 7.7. Peak Acceleration for Ballistic and Guided Entry

2. $(L/D)_1$ and $(L/D)_2$ are the nominal lift-to-drag ratios in the vertical plane defined by eqns 7.54 and 7.45 respectively.
3. $a_{nom}(v)$ and $\dot{h}_{nom}(v)$ are the total acceleration and rate of descent respectively, obtained by simulating a nominal entry without the feedback terms; they are stored as functions of velocity.

Figure 7.7 shows a plot of maximum (lift and drag) acceleration as a function of entry angle. The corresponding figure for the ballistic entry (from Figure 7.6) has been added for comparison. It is evident that, unlike the ballistic entry, there is a range of entry angles (about 7.3 to 9.5 degrees) over which there is little change in the peak acceleration. Within the limits of the permitted extremes of the L/D, *the feedback terms widen the entry corridor* from skip-out at 7.3 degrees to a tolerable maximum acceleration such as 35g at 10.7 degrees. The feedback also renders very low sensitivity to variations of density in the upper atmosphere and uncertainties in the aerodynamic coefficients.

7.5.4 Feedback Guidance in the Lower Atmosphere

The above guidance algorithm has one shortcoming, namely with respect to control of the final position above the surface of the Earth at a low altitude where a parachute might be deployed. The low sensitivity of peak acceleration to entry angle has been noted in our example, but the corresponding sensitivity to down-range distance (about an entry of -8.1 degrees) is 322 km per degree. Using a representative figure of 0.5 degree for the achievable control of entry angle, this implies an uncertainty of 161 km in the final position down-range which would almost certainly be not acceptable. In order to achieve guidance of down and cross-range position it is necessary to compute or estimate in real time the distance-to-go and any cross-range error.

This is determined in the Shuttle guidance system (Ref. [25]) by using approximate expressions when the flight path angle is small. For example, the down-range distance-to-go R_d satisfies

$$\dot{R}_d = v \cos \gamma \quad (7.60)$$

and the velocity equation is approximately

$$\dot{v} = -D - g \sin \gamma \quad (7.61)$$

for drag deceleration D . Therefore the down-range distance for small values of γ is approximately

$$R_d = - \int \frac{vdv}{D} \quad (7.62)$$

The drag D is eliminated by postulating a small constant glide path. When $\dot{\gamma}$ is zero and γ is small, eqn 7.39 provides the approximation

$$D = \frac{g(v^2/v_c^2 - 1)}{L/D} \quad (7.63)$$

For only modest roll manoeuvres (in the case of the Shuttle) L/D could be set at a constant value, thus permitting integration of the down-range distance.

Apollo used large roll manoeuvres even in the lower atmospheres when the above approximation would not be valid. The prediction of down-range distance was handled (Ref. [22]) by computing partial derivatives (on a nominal trajectory) of down-range with respect to acceleration and altitude rate. This approach is generalised below by means of *adjoint functions* with the added benefit that a guidance algorithm is obtained in addition to the predicted estimates of down and cross-range position.

The equations of motion consist of six first order differential equations, time being the independent variable. However time does not appear on the right hand side and, in any case, is a poor choice as the independent variable when storing the variables of a nominal trajectory. Let the velocity at the start of the second guidance phase in the lower atmosphere be v_0 . Define a new independent variable as

$$s = v_0 - v \quad (7.64)$$

which can be assumed to be monotonic increasing from zero to a specified value s_f , corresponding to a final given very low velocity. By dividing both sides of the state equations by the equation in \dot{v} , we obtain five first order differential equations with respect to the independent variable s . Generate a nominal descent trajectory in the lower atmosphere and store the five dependent variables as functions of s . Let the state vector be $(x_i, i = 1, 2..5)$, corresponding to $(\tau, \delta, r, \chi, \gamma)$, but define perturbations about the nominal trajectory to be $(y_i, i = 1, 2..5)$.

By taking a first order expansion of the five state equations, we can obtain a set of five *linear* state equations in the perturbation variables y_i in the form

$$\frac{dy_i}{ds} - \sum_{j=1}^5 a_{ij}y_j = b_i u_1 + c_i u_2 ; \quad i = 1, 2..5 \quad (7.65)$$

The coefficients a_{ij} would be obtained by interpolation from values stored as a function of the independent variable s . They are partial derivatives of the right hand sides of the state equations, calculated analytically or by numerical perturbations. The terms u_1 and u_2 are the small changes in the commands for roll angle as defined in eqns 7.35 and 7.36; they are the guidance or control variables for this phase. The coefficients b_i and c_i are also calculated on the nominal trajectory and stored as a function of s (velocity). We now introduce the adjoint variables λ_i ($i=1,2..5$) which satisfy the linear equations

$$\frac{d\lambda_i}{ds} + \sum_{j=1}^5 a_{ji}\lambda_j = 0 ; \quad i = 1, 2..5 \quad (7.66)$$

Multiply eqn 7.65 by λ_i , eqn 7.66 by y_i and sum over i to yield

$$\frac{d}{ds} \sum_{i=1}^5 \lambda_i y_i - \sum_{i=1}^5 \sum_{j=1}^5 (a_{ij} \lambda_i y_j - a_{ji} \lambda_j y_i) = \sum_{i=1}^5 \lambda_i (b_i u_1 + c_i u_2) \quad (7.67)$$

The double summation reduces to zero, hence

$$\frac{d}{ds} \sum_{i=1}^5 \lambda_i y_i = \sum_{i=1}^5 \lambda_i (b_i u_1 + c_i u_2) \quad (7.68)$$

Integrate this equation from s to s_f with the result

$$\left[\sum_{i=1}^5 \lambda_i y_i \right]_{s_f} = \left[\sum_{i=1}^5 \lambda_i y_i \right]_s + \int_s^{s_f} \sum_{i=1}^5 \lambda_i (b_i u_1 + c_i u_2) ds \quad (7.69)$$

The adjoint variables are not yet defined completely because equations 7.69 are differential equations. To fix the constants of the latter we integrate backwards from s_f to s_k the five adjoint differential equations using interpolation between stored values of the coefficients a_{ij} . This is done twice as follows for a sequence of equally spaced values s_k in the range $(0, s_f)$.

1. Set all $\lambda_i(s_f)$ equal to zero except $\lambda_1(s_f) = 1$. Let the final values of this integration be $\lambda_i^1(s_k)$, $i = 1, 2, 5$.
2. Set all $\lambda_i(s_f)$ equal to zero except $\lambda_2(s_f) = 1$. Let the final values of this integration be $\lambda_i^2(s_k)$, $i = 1, 2, 5$.

For any one of the twofold sets at s_k so generated substitute into eqn 7.69, bearing in mind that $y_1(s_f)$ and $y_2(s_f)$ correspond to perturbations in the final longitude τ and latitude δ . Thus, treating u_1 and u_2 as constants in the interval (s_k, s_f) ,

$$\Delta\tau(s_f) = \sum_{i=1}^5 \lambda_i^1(s_k) y_i(s_k) + B_1 u_1 + C_1 u_2 \quad (7.70)$$

$$\Delta\delta(s_f) = \sum_{i=1}^5 \lambda_i^2(s_k) y_i(s_k) + B_2 u_1 + C_2 u_2 \quad (7.71)$$

where for $j=1$ and 2

$$B_j = \int_{s_k}^{s_f} \left[\sum_{i=1}^5 \lambda_i^j b_i \right] ds \quad (7.72)$$

$$C_j = \int_{s_k}^{s_f} \left[\sum_{i=1}^5 \lambda_i^j c_i \right] ds \quad (7.73)$$

Put eqns 7.70 and 7.71 in matrix notation and specify that the final perturbations (errors) in τ and δ are to be zero, whereupon we have the result

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = - \begin{bmatrix} B_1 C_1 \\ B_2 C_2 \end{bmatrix}^{-1} L y(s_k) \quad (7.74)$$

where the 2x5 matrix L has $(\lambda_1^1, \dots, \lambda_5^1)$ as the first row and $(\lambda_1^2, \dots, \lambda_5^2)$ as the second row, ie. all evaluated at s_k . *Equation 7.74 is a feedback algorithm to be applied and updated at regular intervals of velocity to adjust u_1 and u_2 by means of the roll angle (eqn 7.34) to zero the final errors in longitude and latitude.* The L, B, and C matrices could be pre-computed on the nominal trajectory and stored at a suitable number of subintervals. The adjoint variables are generated to be the partial derivatives of final longitude and latitude with respect to the five state variables (excluding velocity) at points along the nominal trajectory. Recall however that the latter is defined as a function of velocity and not time.

7.6 Software

The simulations of this section were carried out by means of C++ programs grouped together in the file ATM.CPP. Linking to the previously introduced subroutines in ORBIT.CPP and MATRIX.CPP is necessary but routines in those files are not listed below.

The routines included in the file ATM.CPP are as follows.

Main : main program to illustrate atmospheric entry of an unguided capsule and guidance of a vehicle with some lift.

entry : depending on an option number, this routine is used to control

1. simulation of ballistic entry.
2. generation of a nominal trajectory for guided entry using some lift.
3. simulation of feedback guidance about a nominal entry trajectory.

plot_xy : screen plotting routine.

translate : copying components of the state vector into variables used in RHS, etc., or vice versa.

Runge_Kutta4 : Fourth order Runge Kutta numerical integration.

RHS : formation of the right hand sides of the six first order differential equations 7.1 to 7.6.

atmosphere : US Standard Atmosphere 1962.

Earth_radius : Earth radius as a function of latitude.

gravity : gravitational force as a function of latitude and radius.

LD1 : vertical component of L/D ratio for nominal trajectory in lower atmosphere (section 7.5.2).

LD2 : vertical component of L/D ratio for nominal trajectory in upper atmosphere (section 7.5.2).

REFERENCES

1. S M Archer and D D Swooder. Selection of the guidance variable for a re-entry vehicle. *Journal of Guidance and Control*, 2(2):130–138, 1979.
2. A Atzei et al. Rosetta/CNSR-ESA's planetary corenertstone mission. *ESA Bulletin*, (59):19–29, 1989.
3. R H Battin. *An Introduction to the Mathematics and Methods of Astrodynamics*. AIAA Education Series, 1987.
4. R H Battin and R M Vaughan. An elegant Lambert algorithm. *J. of Guidance, Control, and Dynamics*, 7(6):662–670, 1984.
5. M Bello-Mora et al. Near comet orbit planning and navigation tool. *Final Report by GMV of ESA Contract 11250/94/D/IM*, 1997.
6. G J Bierman. *Factorization Methods for Discrete Sequential Estimation*. Academic Press, 1977.
7. R G Brusch. Trajectory optimization for the Atlas/Centaur launch vehicle. *J of Spacecraft*, 14(9):550–555, 1977.
8. A E Bryson and Y C Ho. *Applied Optimal Control*. Blaisdell, 1969.
9. R J Cesareone. A gravity assist primer. *AIAA Student Journal*, 27(2):16–22, 1989.
10. D R Chapman. An approximate analytical method for studying entry into planetary atmospheres. *NASA Technical Report*, (R-11), 1959.
11. J W Cornelisse, H F R Schoyer, and K F Wakker. *Rocket Propulsion and Spaceflight Dynamics*. Pitman, 1979.
12. E J Cramer and J E Bradt. Nlp re-entry guidance: Developing a strtegy for low l/d vehicles. *AIAA Paper*, (88-4123-CP):601–608, 1988.
13. A L Crosbie. *Aerothermodynamics and Planetary Entry: vol. 77, chapter 3*. AIAA, 1981.
14. R S Crowder and J D Moote. Apollo entry aerodynamics. *Journal of Spacecraft*, 6(3):302–307, 1969.
15. L A D'Amario et al. Galileo 1989 VEEGA trajectory design. *Journal of the Astronautical Sciences*, 37(3):281–306, 1989.
16. J K Davies. *Space Exploration*. Chambers, 1992.
17. L C W Dixon et al. Low thrust orbit optimization for interplanetary missions. *Final Report of ESA Study Contract 4774/81*, 1983.
18. K A Ehricke. Instrumented Comets-Astronautics of solar and planetary probes. *Proc. 8th International Astronomical Congress*, 1957.
19. P J Enright et al. Optimal finite thrust spacecraft trajectories using collocation and nonlinear programming. *Journal of Guidance, Control and Dynamics*, 14(5):981–985, 1991.
20. G F Franklin and J D Powell. *Digital Control of Dynamic Sysytems*. Addison-Wesley, 1980.
21. B Friedland. Treatment of bias in recursive filtering. *IEEE Trans. Automatic Control*, AC-14(4):359–367, 1969.

22. A G Graves and J C Harpold. Re-entry targeting philosophy and flight results from apollo 10 and 11. *Proc. AIAA 8th Aerospace Sciences Meeting*, (70-28):1–11, 1970.
23. J W Griffin et al. Shuttle rendezvous radar performance evaluation and simulation. *IEEE Conference on Position Location and Navigation Symposium*, (88CH2675-7):236–245, 1988.
24. W Haeussermann. Saturn launch vehicle's navigation guidance, and control system. *Automatica*, 7:537–556, 1971.
25. J C Harpold and A G Graves. Shuttle entry guidance. *Journal of the Astronautical Sciences*, 27(3):239–268, 1979.
26. M Hechler and G Schwehm. Rosetta: ESA's Cornerstone Mission. *ESA Bulletin*, (77):7–18, 1994.
27. S Jallade et al. Trajectory optimization and guidance strategies for re-entry. *ESA Special Publication*, (SP-323):137–144, 1991.
28. R E Kalman and R S Bucy. New results in linear filtering and prediction theory. *Trans. ASME, Series D, J of Basic Eng.*, 83:95–108, 1961.
29. M H Kaplan. *Modern Spacecraft Dynamics and Control*. Wiley, 1976.
30. A R Klumpp. Rendezvous guidance trajectory shaping. *IEEE Aerospace and Electronics Systems Magazine*, 2(2):17–22, 1987.
31. D F Lawden. Interplanetary rocket trajectories. *Advances in Space Science, Academic Press*, 1:1–53, 1959.
32. C L Lawson and R J Hanson. *Solving Least Squares Problems*. Prentice-Hall, 1974.
33. S M Lichten. Estimation and filtering for high-precision gps positioning applications. *Manuscripta Geodaetica*, 15:159–176, 1990.
34. R G Marsden. Ulysses explores the south pole of the sun. *ESA Bulletin*, (82):48–56, 1995.
35. A R Martin. Nuclear propulsion of spacecraft. *Journal of British Interplanetary Society*, 48(12):517–524, 1995.
36. T D Moyer. Mathematical formulation of the double precision orbit determination program. *JPL Technical Report*, 32(1527), 1971.
37. W G Naumann. Ariane-4:Europe's launcher for the next decade. *ESA Bulletin*, (55):75–83, 1988.
38. V Neal. *Spaceflight, a Smithsonian Guide*. Macmillan, 1995.
39. Authors not specified. *Planetary and Lunar Coordinates*. HM Stationary Office and US Gov. Printing Office, 1983.
40. M Noton. *Modern Control Engineering*. Pergamon Press, 1972.
41. M Noton. Orbit strategies and navigation near a comet. *ESA Journal*, 16:349–362, 1992.
42. M Noton and S V Salehi. Low-thrust navigation to comets or asteroids. *ESA Journal*, 11:215–231, 1987.
43. J M Pairot. Automatic rendezvous in space: a new European technology. *Navigation*, 41(164):519–529, 1993.
44. J J Pocha. *An introduction to mission design for Geostationary Satellites*. Kluwer, 1987.
45. F J Regan. *Re-entry Vehicle Dynamics*. AIAA, 1984.
46. G Saccoccia. European electric propulsion activities in the era of application. *Journal of British Interplanetary Society*, 48(12):487–500, 1995.
47. J A Saponaro and S L Copps. Operations and functions of the Apollo guidance computer during rendezvous. *3rd IFAC Symposium on Automatic Control in Space*, (2):52–56, 1970.

48. J B Serrano-Martinez. Use of simulation tools and facilities for rendezvous and docking missions. *AGARD Conference on Space system design and development testing*, CP-561(17):1–12, 1995.
49. S K Shrivastava. Orbital perturbations and station-keeping of communication satellites. *J. of Spacecraft and Rockets*, 15(2):67–78, 1978.
50. E M Standish. Planetary ephemerides. *Astronomy and Astrophysics*, 114:297–302, 1982.
51. G Strang. *Introduction to Linear Algebra*. Wellesley-Cambridge, 1993.
52. E Stuhlinger et al. Comet nucleus sample return mission with electrically propelled spacecraft. *Journal of British Interplanetary Society*, 39(6):273–281, 1986.
53. C W Uphoff. The art and science of lunar gravity assist. *AAS Paper, Proc. of AAS/NASA International Symposium*, 89(170):333–346, 1989.
54. W S Vaning and J F Fay. Effect of lunar gravity assist on interplanetary trajectories. *AAS Paper*, 91(477):490–506, 1991.
55. M D Waterman. Guidance law for the Ariane On-board computer. *J of the British Interplanetary Society*, 29:402–407, 1976.
56. K H Well and S R Tandon. Rocket ascent trajectory optimization via recursive quadratic programming. *J of the Astronautical Sciences*, 30(2):101–116, 1982.
57. J R Wertz. *Spacecraft Attitude Determination and Control*. Reidel, 1980.
58. J R Wertz and W J Larson. *Space Mission Analysis and Design*. Kluwer, 1990.
59. S N Williams and J M Longuski. Automated design of multiple encounter gravity-assist trajectories. *Proceedings of the AIAA/AAS Astrodynamics Conference*, 2:985–994, 1990.
60. R J Wright and J V Sponnick. A ring laser gyro based navigator for space launch vehicle guidance. *IEEE Aerospace and Electronics Systems Magazine*, 4(3):29–38, 1989.
61. K E Wurster. Lifting entry vehicle mass reduction through integrated thermostructural trajectory design. *Journal of Spacecraft*, 20(6):589–596, 1983.
62. E V Zoby et al. Approximate convective heating equations for hypersonic flows. *Journal of Spacecraft*, 18(1):64–70, 1981.

APPENDICES

A.1 Perturbations from the Sun and Moon in Geostationary Orbit

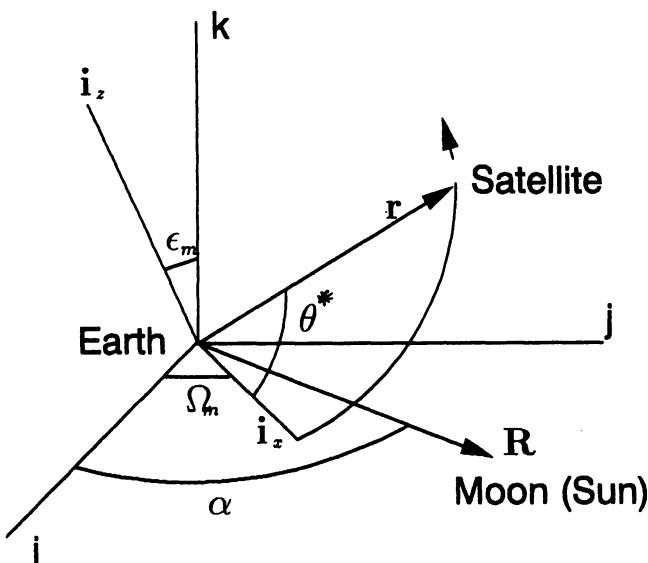


Fig. A.1. Perturbations from the Sun or Moon

Figure A.1 illustrates a geostationary circular orbit with reference to axes such that k is the normal to the orbital plane of the Moon (Sun). The latter is at distance R_m . In the axes $(\mathbf{i}_x, \mathbf{i}_y, \mathbf{i}_z)$ the position vector of the satellite is

$$\mathbf{r} = a(\cos \theta^*, \sin \theta^*, 0) \quad (\text{A.1})$$

where

$$\theta^* = \omega_e t \quad (\text{A.2})$$

in a circular orbit. The position vector of the Moon (Sun) in the same axes is

$$\mathbf{R} = R_m(\cos \alpha, \cos \epsilon_m \sin \alpha, -\sin \epsilon_m \sin \alpha) \quad (\text{A.3})$$

where

$$\alpha = \omega_e t + \phi \quad (\text{A.4})$$

Having introduced the phase angle ϕ , the angle Ω_m of Figure A.1 can be set to zero without loss of generality.

Referring to section 3.4 and in particular to eqn 3.96 it is noted that the perturbation acceleration is

$$\mathbf{f} = f_0(\mathbf{r} \cdot \mathbf{R}) \mathbf{R} / (a R_m^2) \quad (\text{A.5})$$

where

$$f_0 = 3\nu\mu a / R_m^3 \quad (\text{A.6})$$

Thus

$$\mathbf{f} = f_0(\cos \theta^* \cos \alpha + \sin \theta^* \cos \epsilon_m \sin \alpha) \begin{bmatrix} \cos \alpha \\ \cos \epsilon_m \sin \alpha \\ -\sin \epsilon_m \sin \alpha \end{bmatrix} \quad (\text{A.7})$$

Use of the tangential component in conjunction with eqn 3.52 confirms that the mean value of da/dt is zero. In order to avoid the singularity at zero inclination we will employ the results of section 3.2.6 with respect to axes rotated through 90 degrees. In that case inclination of the orbit (original axes) corresponds to changes in both i and Ω in the new axes. Although eqn 3.53 results in a zero mean value for di/dt , this is not true for $d\Omega/dt$. Denote the latter in new axes as $dinc/dt$, then by substituting the z-component of eqn A.7 into eqn 3.54

$$\frac{dinc}{dt} = -\frac{af_0 \sin \theta^*}{h} (\cos \theta^* \cos \alpha + \sin \theta^* \cos \epsilon_m \sin \alpha) \sin \epsilon_m \sin \alpha \quad (\text{A.8})$$

Substitute for f_0 and take average values first over the satellite orbit and then over one orbit of the Moon (Sun).

$$\left[\frac{dinc}{dt} \right]_{av} = -\frac{3\nu\mu \sin 2\epsilon_m}{8\omega_e R_m^3} \quad (\text{A.9})$$

A.2 Perturbations due to Solar Pressure in Geostationary Orbit

The acceleration due to solar pressure is

$$\mathbf{f} = -(f_s/R_s) \mathbf{R} \quad (\text{A.10})$$

where \mathbf{R} is the radius vector of length R_s from Earth to Sun, and f_s is given by eqn 3.119. From eqn A.3 and Figure A.1 the above acceleration is therefore

$$\mathbf{f} = f_s(-\cos \alpha, -\cos \epsilon_s \sin \alpha, \sin \epsilon_s \sin \alpha) \quad (\text{A.11})$$

in the $(\mathbf{i}_x, \mathbf{i}_y, \mathbf{i}_z)$ of that figure. Subscript s denotes the Sun. In the same axes

$$\mathbf{i}_r = (\cos \theta^*, \sin \theta^*, 0) \quad (\text{A.12})$$

and

$$\mathbf{i}_\theta = (-\sin \theta^*, \cos \theta^*, 0) \quad (\text{A.13})$$

Therefore, continuing to use the notation of section A.1

$$f_r = -f_n = -f_s(\cos \theta^* \cos \alpha + \sin \theta^* \cos \epsilon_s \sin \alpha) \quad (\text{A.14})$$

$$f_\theta = f_t = f_s(\sin \theta^* \cos \alpha - \cos \theta^* \cos \epsilon_s \sin \alpha) \quad (\text{A.15})$$

Substitute these expressions in eqns 3.50 and 3.51, assuming rotated axes with $i = \pi/2$.

$$\frac{de_x}{dt} = f_s [\cos \theta^* \sin \theta^* \cos \alpha - (\cos^2 \theta^* + 1) \cos \epsilon_s \sin \alpha] / v \quad (\text{A.16})$$

$$\frac{de_y}{dt} = f_s [-\cos \theta^* \sin \theta^* \cos \epsilon_s \sin \alpha + (\sin^2 \theta^* + 1) \cos \alpha] / v \quad (\text{A.17})$$

Take mean values over one orbit to yield

$$\left[\frac{de_x}{dt} \right]_{av} = -\frac{3f_s}{2v} \cos \epsilon_s \sin \alpha \quad (\text{A.18})$$

$$\left[\frac{de_y}{dt} \right]_{av} = +\frac{3f_s}{2v} \cos \alpha \quad (\text{A.19})$$

where

$$\alpha = \omega_s t + \phi \quad (\text{A.20})$$

Approximate $\cos \epsilon_s$ equal to unity and integrate, setting the constant such that initial values are zero. After some re-arrangement we obtain the result

$$\begin{aligned} e^2 &= e_x^2 + e_y^2 \\ &= \left[\frac{3f_s}{\omega_s v} \right]^2 \sin^2 \frac{\omega_s t}{2} \end{aligned} \quad (\text{A.21})$$

When the time is much less than one year the sin function can be approximated by the angle and consequently

$$e \simeq \frac{3f_s t}{2v} \quad (\text{A.22})$$

ie. the rate of increase of eccentricity is approximately equal to $1.5f_s/v$.

A.3 Angular Information from Doppler Tracking

Referring to the analysis of section 4.2 and Figure 4.1, the measured range at a ground tracking site is

$$\begin{aligned} r_m &= \sqrt{d^2 + s^2 - 2d.s} \\ &\simeq d[1 - (s/d)\mathbf{d}.s/s + s^2/d^2] \end{aligned} \quad (\text{A.23})$$

By differentiation, the measured range-rate is

$$\dot{r}_m = (d/r_m) \left[\dot{d} - \mathbf{d}.\dot{s}/d + (s/d)(\dot{s} - \mathbf{d}.s/s) \right] \quad (\text{A.24})$$

where \mathbf{s} and $\dot{\mathbf{s}}$ are given by eqns 4.3 and 4.4 respectively. Note that the angle α of eqn 4.2 is associated with sinusoidal terms at the frequency of the Earth's sidereal spin rate. Put

$$\mathbf{d} = d(\cos D \cos RA, \sin D \sin RA, \sin D) \quad (\text{A.25})$$

i.e. introducing the Right Ascension and Declination of the object in Earth-based Equatorial axes. Combine now eqns A.23, A.24 and A.25 but (at great distances from the Earth) ignore terms the order of s/d . The result is

$$\begin{aligned} \dot{r}_m &\simeq \dot{d} - \mathbf{d}.\dot{s} \\ &= \dot{d} - \omega_e r_0 \cos \theta_0 \cos D \sin(\phi_0 \omega_e t + \phi'_0 - RA) \end{aligned} \quad (\text{A.26})$$

Equation A.26 reveals that when tracking a spacecraft at great distances from the Earth, the doppler signal will consist of a mean value (range-rate relative to the centre of the Earth) modulated at a frequency equal to the Earth's sidereal spin rate ω_e . The amplitude and phase of this sinusoidal signal gives information about the Declination and Right Ascension of the spacecraft. This expression is not useful if the Declination is nearly zero but most space missions are near the Ecliptic plane, (with a Declination of 23.45 degrees) so the result fails only when the spacecraft is approximately along the direction of the Spring or Autumn Equinoxes.

A.4 Ephemeris of the Sun

The ephemeris of the Sun relative to Earth, or equivalently the Earth relative to the Sun, can be generated from the known orbital elements of the Earth (Ref. [39]). However, since the eccentricity is so small a convenient formula is available (Ref. [57]) which is sufficiently accurate for many purposes.

In the following expressions for the ephemeris of the Sun all angles are in degrees. The mean longitude is calculated by

$$L_m = 279.696678 + 0.9856473354d \quad (\text{A.27})$$

The mean anomaly is

$$M = 358.475845 + 0.985600267d \quad (\text{A.28})$$

where d is the number of Julian days minus 2415020. A correction term is

$$\delta L = 1.918 \sin M + 0.02 \sin 2M \quad (\text{A.29})$$

which is employed to give the true longitude as

$$L = L_m + \delta L \quad (\text{A.30})$$

and the true anomaly as

$$\theta = L_m + \delta L \quad (\text{A.31})$$

The distance to the Sun is then given by

$$R = AU(1 - e^2)/(1 + e \cos \theta) \quad (\text{A.32})$$

where AU is the value of the astronomical unit and $e = 0.016751$. Thus the position and velocity vectors (in Ecliptic axes) of the Earth relative to the Sun are given by

$$\mathbf{R} = \begin{bmatrix} -R \cos L \\ -R \sin L \\ 0 \end{bmatrix} \quad (\text{A.33})$$

$$\mathbf{V} = \begin{bmatrix} \dot{L}R \sin L - \dot{R} \cos L \\ -\dot{L}R \cos L - \dot{R} \sin L \\ 0 \end{bmatrix} \quad (\text{A.34})$$

where

$$\dot{R} = \frac{Re \sin \theta}{(1 + e \cos \theta)}(\dot{M} + \delta \dot{L}) \quad (\text{A.35})$$

The derivatives with respect to time (in days) are obtained from the above equations but remember that eqn A.35 requires \dot{L} , \dot{M} and $\delta \dot{L}$ in radians per day.

A.5 Sequential UD Covariance Filter

Linear Time Update. It is assumed that the equations for the state vector x of n components are as follows.

$$x(k+1) = A_x(k)x(k) + C(k)w(k) \quad (\text{A.36})$$

where $w(k)$ is a vector of n_w independent random inputs of zero mean value and unit variance. The time update (mapping) is achieved in the UD factorized form of the Kalman Filter as follows; for the proof and more details see Bierman's section 6.4, Ref. [6].

$$\tilde{x}(j+1) = A_x(j)\hat{x}(j) \quad (\text{A.37})$$

where $\hat{x}(j)$ is the best estimate at instant j . The covariance matrix before the time update is \hat{P} such that

$$\hat{P} = \hat{U}\hat{D}\hat{U}^T \quad (\text{A.38})$$

where \hat{U} is $(n \times n)$ upper triangular and \hat{D} is $(n \times n)$ diagonal. Assume that \hat{U} and \hat{D} are available, then the updated \tilde{U} and \tilde{D} are calculated by the following Modified Weighted Gram-Schmidt recursive algorithm. First put

$$S = [A_x \hat{U} \quad C] \quad (\text{A.39})$$

and

$$D = \text{Diag}(\hat{D}, I) \quad (\text{A.40})$$

assuming unit process noise input. Let s_i^T denote the i th row of the n by $(n + n_w)$ matrix S . Starting from

$$v_k = s_k \quad k = 1, \dots, n \quad (\text{A.41})$$

carry out the following recursions for $j = n - 1, \dots, 1$.

$$\tilde{D}_{j+1} = v_{j+1}^T D v_{j+1} \quad (\text{A.42})$$

$$\tilde{U}(k, j+1) = v_k^T D v_{j+1} / \tilde{D}_{j+1}; \quad k = 1, \dots, j \quad (\text{A.43})$$

Replace v_k by

$$v_k - \tilde{U}(k, j+1)v_{j+1} \quad k = 1, \dots, j \quad (\text{A.44})$$

and

$$\tilde{D}_1 = v_1^T D v_1 \quad (\text{A.45})$$

In this way the time update is achieved of \hat{U} and \hat{D} to \tilde{U} and \tilde{D} , the update of the state vector being given by A.37 above.

Linear Measurement Update. Consider now the update, given \tilde{U} and \tilde{D} , when one scalar measurement z becomes available as follows.

$$z = h^T x + \nu \quad (\text{A.46})$$

where the measurement noise ν has zero mean and variance r . The calculation of \hat{U} and \hat{D} (ie. after the measurements) is obtained by the following algorithm; see Bierman's section 5.3 [6].

$$f = \tilde{U}^T h \quad (\text{A.47})$$

$$v = \tilde{D}f \quad (\text{A.48})$$

Initiate

$$\hat{d}_1 = \tilde{d}_1 r / \alpha_1, \quad \alpha_1 = r + v_1 f_1, \quad K_2^T = (v_1, 0, \dots, 0) \quad (\text{A.49})$$

and use \tilde{u}_j to denote the j 'th column vector of \tilde{U} . Cycle $j = 2, 3, \dots, n$

$$\alpha_j = \alpha_{j-1} + v_i f_j \quad (\text{A.50})$$

$$\hat{d}_j = \tilde{d}_j \alpha_{j-1} / \alpha_j \quad (\text{A.51})$$

$$\hat{u}_j = \tilde{u}_j - f_j K_j / \alpha_{j-1} \quad (\text{A.52})$$

$$K_{j+1} = K_j + v_j \tilde{u}_j \quad (\text{A.53})$$

Then the Kalman gain is

$$K = K_{n+1} / \alpha_n \quad (\text{A.54})$$

for the measurement update

$$\hat{x} = \tilde{x} + K(z - h^T \tilde{x}) \quad (\text{A.55})$$

A.6 Global Positioning System

The Global Positioning System (GPS, Ref. [52]) consists of at least 21 satellites launched by the United States Department of Defense equally spaced in six orbit planes at about 20,000 km altitude. Although it was designed primarily for U.S. military applications it has been made available internationally and is in use for a range of civil and commercial applications, on the ground, at sea, and in the air. It can be used to determine the position of Earth satellites provided that their altitude is well below 20,000km and they carry the necessary GPS receiver.

The basic operation of GPS with respect to satellites is explained first in two dimensions by reference to Figure A.2. The GPS satellites transmit signals continuously at L-band (about 1.5 GHz) modulated in a binary manner at 10 Mbits/sec by a pseudo random sequence, ie. it appears to be random but it actually repeats itself after a suitable interval. The sequences of the GPS satellites are synchronized very accurately by means of a Caesium clock. When a spacecraft receives such signals from two GPS satellites *the time difference between the two sequences can be determined*. Thus if the time difference is Δt_{12} , then referring to Figure A.2,

$$R_2 - R_1 = c\Delta t_{12} \quad (\text{A.56})$$

where c is the velocity of radio propagation (light). Note that because the receiving satellite does not have the GPS clock time, the distances R_1 and R_2 are not available as direct measurements. As a result of a known fundamental property of a hyperbola, equation A.56 implies that the spacecraft lies on a hyperbola as shown in the figure. Assume now that signals from GPS satellites 2 and 3 are also processed to determine

$$R_3 - R_2 = c\Delta t_{23} \quad (\text{A.57})$$

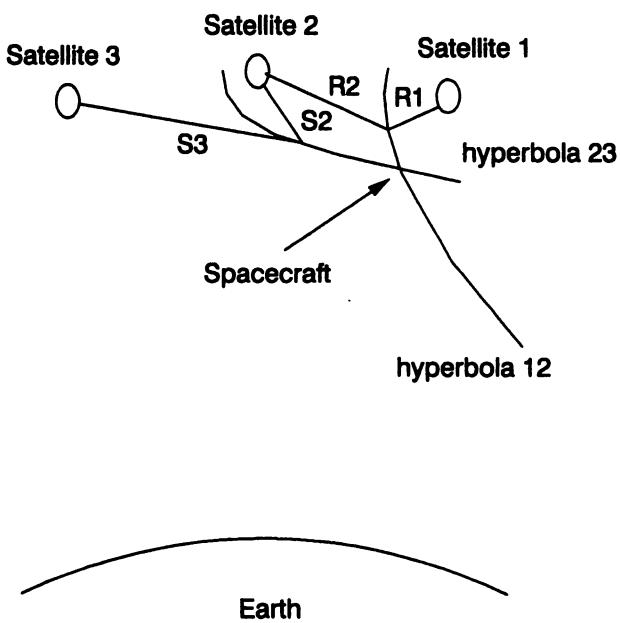


Fig. A.2. GPS in Two Dimensions

which defines a second hyperbola. The two hyperbola intersect in two points but the spacecraft position can be uniquely fixed because one intersection (not shown on the diagram) will be at a high altitude above the GPS satellites.

In order to extend the discussion to three dimensions, assume that signals are received from four satellites. The locus of possible points, given the measured time difference Δt_{12} , becomes a surface generated by rotating the hyperbola about the line joining GPS satellites 1 and 2, and similarly for two other measured time differences. *The spacecraft can consequently be located in three dimensions at the intersection of three hyperbolic surfaces.* In practice three simultaneous second-order equations must be solved (eg. iteratively given a first estimate) but such calculations can be embraced within a SRIF or equivalent updating algorithm (Ref. [33]).

The GPS system operates two codes, (a) the coarse-acquisition C/A code for approximate positioning, and (b) the P code for refined purposes. Although the P code is available only to U.S. officially approved users, even the C/A code allows the determination of position to an accuracy of 50-100 m. The P code yields an accuracy of 10-15 m and a sixfold improvement is possible by *differential GPS tracking* which exploits additional ground-based receivers (Ref. [33]).

INDEX

- ablation, 145
- acceleration, 147
- accelerometers, 53
- adjoint variables, 162
- aerodynamic forces, 39
- aerodynamics, 143
- alpha, 143
- angle of attack, 44
- aphelion, 132
- apocentre, 3
- apogee, 3
- Apollo, 141, 143, 147, 158
- argument of pericentre, 5
- Ariane, 34, 44
- Aries, 5
- ascending node, 5
- ascent, 41, 54
- astronomical unit, 8
- asymptote, 118
- atmosphere model, 144
- atmosphere, 141, 40
- atmospheric entry, 145, 164
- ballistic coefficient, 67, 144
- ballistic entry, 152
- ballistic coefficient, 148
- bias terms, 96
- blunt nose, 143
- burn time, 45
- calculus of variations, 47
- camera, 104, 104, 106
- celestial axes, 5
- characteristic area, 143
- chemical propulsion, 117
- circular velocity, 148
- closest approach, 119
- coast arc, 129
- comet, 104
- Comet, 122
- comets, 125
- cone, 143
- conic, 3
- consider matrices, 96
- consider matrix, 99
- constraints, 48
- control vector, 137
- control, 156
- controllability, 133, 134
- correction, 153
- corridors, 145, 151
- costates, 48, 127
- covariance filters, 100
- covariance matrix, 92, 95, 112
- Cowell, 25
- cross-range, 161
- deceleration, 144
- Declination, 112
- declination, 5
- deflection, 119, 121
- density, 144
- descent rate, 157, 159
- docking, 79
- doppler, 88, 112
- down-range, 153, 161
- drag coefficient, 144
- drag deceleration, 157, 159
- drag, 40, 143, 155
- Earth gravity assist, 122
- Earth radius, 38, 82
- east-west perturbations, 71, 72
- eccentric anomaly, 10
- eccentricity vector, 60
- eccentricity, 3, 7
- ecliptic axes, 5
- electric propulsion, 31, 125, 130
- emissivity, 145
- Encke, 26
- end point guidance, 137, 163
- Energia, 34
- energy, 4, 117
- entry angle, 147, 153, 161

- entry guidance, 156
- Ephemeris of the Sun, 112
- equatorial axes, 5
- equatorial radius, 142
- errors of execution, 78
- errors of navigation, 78
- Euler angles, 5
- execution errors, 113
- feedback guidance, 137, 158, 161
- final constraints, 133
- final dispersion, 114
- final guidance, 163
- final mass, 130
- final position, 161
- fixed time guidance, 112
- flat Earth, 43
- flight equations, 141
- flyby, 23
- gamma, 141
- geostationary orbit, 109, 110
- geostationary orbits, 69
- geostationary transfer orbit, 101
- geostationary, 42
- GPS, 75, 83, 115
- gravitational data, 8
- gravitational field, 62
- gravitational forces, 38
- gravitational potential, 62
- gravity assist, 109, 116
- gravity turn, 43
- ground tracking, 88, 102, 104, 153
- guidance, 51, 52
- guided entry, 155
- gyros, 53
- Hale-Bopp, 88
- Hamiltonian, 127
- heating, 145, 147, 151
- Hohmann transfer, 17
- horizontal lift, 156
- Householder, 94, 99
- hyperbolic anomaly, 11
- hyperbolic escape, 19, 110
- inclination, 5
- independent variable, 162
- inertial navigation, 53, 156, 159
- information matrix, 93
- injection errors, 111, 115
- injection, 18, 111, 132
- Julian days, 21
- Jupiter, 22, 23, 110, 115, 119
- Kalman Filter, 100
- Kepler, 82
- L/D, 148
- Lambert's algorithm, 76, 122
- Lambert's theorem, 12
- Laplace, 85
- launch geometry, 18
- launch systems, 33
- launch, 132
- least-squares, 90, 92
- level flight, 148
- lift, 40, 143, 155
- lift-to-drag ratio, 148, 158
- linear quadratic control, 136
- linear state equations, 136
- local linearization, 91
- low thrust guidance, 137
- low thrust, 125, 126
- LQ algorithm, 159
- manned vehicle, 147
- Mars, 144
- Maximum Principle, 128
- mean anomaly, 10
- measurement errors, 102, 106
- measurement update, 100
- measurements, 91
- midcourse corrections, 110
- midcourse guidance, 115
- midcourse, 109
- miss components, 114
- Moon, 109
- navigation, 52, 156
- NEP, 125
- Newton, 82
- nominal trajectory, 156
- non-constant parameters, 97
- non-spherical Earth, 62
- nonlinear system, 95
- normalized errors, 92
- north-south perturbations, 70
- nose radius, 147
- nozzle, 30, 39
- nuclear propulsion, 125
- numerical integration, 25
- oblate Earth, 63
- offset, 145, 155
- optimal control, 47, 128, 136
- optimization, 126, 129, 133, 159
- optimum ascent, 42, 44
- orbit determination, 81, 153
- orbit plane, 5
- orbital elements, 4, 7, 41, 89
- orbital estimate, 89
- orbital estimates, 83, 85
- orbital manoeuvres, 16
- orbital plane change, 19

- osculating elements, 129
- osculating orbit, 55
- outer planets, 117, 122
- parachute, 161
- parallax, 86
- parameter estimation, 91
- parking orbit, 110
- partial derivatives, 91, 92, 96
- patched conics, 20, 122
- payload ratio, 33
- payload, 31
- periapsis, 3
- pericentre, 3
- perigee, 3
- perturbations from drag, 67
- perturbations from oblateness, 64
- perturbations from Sun and Moon, 66
- perturbations, 109, 110
- planetary data, 8
- planetary encounters, 116
- planetary missions, 109
- pointing angles, 130, 133
- Pontryagin, 128
- position and velocity, 6
- propellant requirements, 123
- propellant, 45
- propellants, 31
- propulsion requirements, 116
- propulsion, 30
- Proton, 34
- radiation, 145
- random inputs, 136
- random walk, 98
- range, 88, 104
- range-rate, 88, 104
- recovery, 153
- reference area, 143
- refinement of orbits, 90
- rendezvous, 75, 79, 127, 129
- residuals, 92
- Right Ascension, 112
- right ascension, 5
- rocket equation, 31
- rocket thrust, 39
- rocket, 30, 33
- roll angle, 155
- roll, 155
- rolling, 158
- Rosetta, 105, 141
- rotated axes, 132
- RST axes, 113
- Runge-Kutta, 25
- Saturn, 34
- semi-major axis, 4
- semilatus rectum, 3, 7
- SEP, 125
- separation theorem, 137
- sequential estimation, 97
- shaping trajectory, 75
- Shuttle, 75, 141, 161
- sidereal day, 8
- singular value decomposition, 134
- singular values, 134
- skip out, 153, 158
- software, 27, 54, 80, 108, 123, 139, 164
- solar day, 8
- solar polar mission, 22
- solar pressure, 110
- solar system, 82
- specific impulse, 30, 39, 125
- sphere of influence, 110
- spheres of influence, 20
- spherical harmonics, 62
- square root matrices, 93
- SRIF, 93, 94, 101, 102, 104
- staging, 31
- stagnation point, 145
- stars, 105
- state equations, 41, 127, 129
- state transition matrix, 112, 132
- static stability, 155
- stationkeeping, 70, 71, 72
- steering law, 49
- stochastic disturbances, 98
- stochastic inputs, 136
- strapdown, 53
- structural efficiency, 31
- structure, 45
- Sun, 109
- swingby, 23, 116, 118
- terminal conditions, 129
- throttling, 133
- thrust, 30, 45
- time in orbit, 9
- time update, 98
- Titan, 34, 141, 144
- total acceleration, 151
- tracking sites, 101
- trajectory, 35
- transfer time, 15
- trim, 155
- true anomaly, 3, 7
- two-body problem, 2
- UD covariance filter, 101, 113
- unguided entry, 152
- unmanned capsule, 147

variable time guidance, 113
variation of elements, 55, 60
vehicle- centred, 35
Venus, 144
vertical lift, 156
Voyager, 117

weather, 153, 159
web sites, 28
wind axes, 36,142
Wirtanen, 105
zero eccentricity, 60
zero inclination, 61