

Lei Chen · Xian-Zong Bai
Yan-Gang Liang · Ke-Bo Li

Orbital Data Applications for Space Objects

Conjunction Assessment and Situation
Analysis

 Springer

Orbital Data Applications for Space Objects

Lei Chen · Xian-Zong Bai
Yan-Gang Liang · Ke-Bo Li

Orbital Data Applications for Space Objects

Conjunction Assessment and Situation
Analysis

 國防工業出版社
National Defense Industry Press

 Springer

Lei Chen
National University of Defense Technology
Changsha
China

Yan-Gang Liang
National University of Defense Technology
Changsha
China

Xian-Zong Bai
National University of Defense Technology
Changsha
China

Ke-Bo Li
National University of Defense Technology
Changsha
China

ISBN 978-981-10-2962-2 ISBN 978-981-10-2963-9 (eBook)
DOI 10.1007/978-981-10-2963-9

Jointly published with National Defense Industry Press, Beijing, China

Library of Congress Control Number: 2016954628

© National Defense Industry Press, Beijing and Springer Nature Singapore Pte Ltd. 2017

This work is subject to copyright. All rights are reserved by the Publishers, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

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

The publishers, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publishers nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer Nature Singapore Pte Ltd.

The registered company address is: 152 Beach Road, #22-06/08 Gateway East, Singapore 189721, Singapore

Preface

Space objects generally refer to all in-orbit spacecrafts, rocket bodies, and space debris, etc. Unlike meteoroid from the natural world, all space objects are made by mankind. Nowadays, more and more space objects are produced due to the frequent space activities performed by human beings, resulting in a greater possibility of space collision. The Chinese government is continuously promoting and forcefully implementing relevant policies and measures for space environment protection and maintenance, just as other large and responsible spacefaring powers have been doing all the time.

However, the continuous increasement of space objects is inevitable, and all mitigation measures taken can only gear down the increase rate of the space object. Conjunction assessment and collision avoidance maneuvers are the topics and countermeasures which ought to be further studied and carried out for effective applications by spacefaring powers. As the conjunction assessment study is based on the space catalog data (including basic information, orbital information, and characteristic information of space object) obtained from space surveillance network, how to take good use of these data and how to do the data mining comprehensively are the key to the strengthening of the conjunction assessment capacity. Furthermore, every space object performs in-orbit operation in strict accordance with orbital dynamics, but the synthetic movement of most space objects is virtually disordered. How to pinpoint the inherent principles of such motions from huge amount of catalog data is a challenge to us.

The space object orbital data, as an important part of space catalog data, can comprehensively be applied to conjunction assessment, orbital anomaly and space event detection, orbital error analysis and precision optimization, space debris environmental and flow analysis. Through computation and analysis with reference to orbital data, we can obtain the spatial position and velocity information of the space object at current time and a period of time in expected future. Thus, the foundation can be built upon it to let us know more about the space object distribution, and a strong support can be given to the conjunction assessment and collision avoidance maneuvers of in-orbit spacecrafts. The space object catalog data accumulated on a long-term basis can provide us the historical changing process

of space object orbits. Analysis of such data can lead to some intrinsic principles behind the orbital data, for example, the spacecraft orbital maneuver events. The innovative researches on such principles will be conducive to the expansion of the depth and width of space situational analysis and improvement of the ability to acquire the orbital characteristics of space objects.

This book discussed the applications of orbital data of space objects in target conjunction assessment and space situational analysis as well as relevant theories and methodologies. There are totally 7 chapters. The origins, types, and application purposes of space object orbital data are introduced in Chap. 1, where the applications of space object orbital data in target conjunction assessments and space situational analyses are emphasized; an overview of the space object orbit prediction error is presented in Chap. 2, the propagation methods of orbit prediction error based on orbit models are given, and so are the propagation methods based on relative dynamic theories; an error analysis method based on historical orbital data is introduced in Chap. 3, the periodic characteristics of the orbit prediction error are presented, and the error fitting is done by bringing in the Poisson series; an introduction to the analytical and numerical methods commonly applied to the space object approaching analysis is unfolded in Chap. 4; derivations of several explicit expressions of the collision probability under circular and ordinary orbital conditions through analysis of approaching distance are presented in Chap. 5, so as to have the collision probability expressed as explicit functions of the approaching distance or approaching geometric relationship, and errors and application scopes of the explicit expressions are also analyzed; the sensitivity of collision probability, maximum collision probability, false dismissal probability and false alarm probability of conjunction assessment, and multi-factor considered general collision risk assessment method on the basis of collision probability explicit expressions are deeply studied in Chap. 6; orbital anomaly and space event detection methods based on historical orbital data and assessments on such methods are illustrated in Chap. 7. Hence, this book can be used as a reference book for engineers and technical staffs specialized in spacecraft dynamics, spacecraft observation and control, space surveillance, space situational awareness, space debris research, and other related fields. It can also be used as a supporting course material for postgraduate students in relevant majors.

It should be honestly pointed out that this book is far from complete and perfect. For author's limited knowledge, there must be some small typos or even theoretical imperfections in this book. If the readers have any comment or suggestion about this book, we are really grateful for your direct contacting with us through chenl@nudt.edu.cn. Your help will be highly appreciated.

Changsha, China
June 2016

Lei Chen
Xian-Zong Bai
Yan-Gang Liang
Ke-Bo Li

Acknowledgements

This book is a joint research product mainly based on two new doctoral dissertations of our research group and also accredited to the researchers of our reference source. This book is attributable to these collaborators. We are forever indebted to our colleagues and students, without whom this book and the research on which it is based could not be carried on.

Firstly, we want to express our gratitude to Prof. Chen Xiaoqian and Prof. Tang Guojin, senior officials of our college, who give great support and care to our research group. Special thanks should go to Mr. Wang Dongya, Mr. Huang Jianyu, and Mr. Zhang Guojun, who give great supports to our research programs. We would also like to pay our gratitude to Prof. Zhou Bozhao and Prof. Zhao You, who provide many enlightening ideas in our academic discussions. We are also grateful to several insightful and diligent students, including Chai Hua, Zhang Taotao, Hao Jia, Cao Yuhui, and Hu Yunpeng, who have made great efforts and done much work during the compilation of this book. Finally, the work in this book is partially supported by the National Natural Science Foundation of PR China (Grant No. 51405499) and National Basic Research Program of China (“973” Project) (Grant No. 2013CB733100).

Contents

1 Introduction	1
1.1 Data Source—Space Surveillance System	3
1.1.1 Space Surveillance	4
1.1.2 Space Surveillance System	5
1.1.3 Space Surveillance Networks	7
1.2 Data Type-Space Object Catalog	9
1.2.1 Basic Information of Space Catalog Data	9
1.2.2 Orbital Model of Catalog Data	10
1.2.3 U.S. Catalog Data	11
1.3 Data Application—Conjunction Assessment and Situational Analysis	13
1.3.1 Conjunction Assessment	13
1.3.2 Situational Analysis	20
References	21
2 Orbital Prediction Error Propagation of Space Objects	23
2.1 Definition of Coordinate System	23
2.2 Overview of Orbital Prediction Errors	25
2.2.1 Category of Orbital Prediction Errors	25
2.2.2 Model-Based Orbital Error Propagation	28
2.2.3 Comparison with High Precision Ephemeris	29
2.3 Covariance Analysis Describing Function Technique	29
2.3.1 Covariance Analysis of Linear System	29
2.3.2 Covariance Analysis of Nonlinear System	31
2.3.3 Gaussian Distribution-Based Describing Function	33
2.3.4 CADET’s Application on Orbit	36
2.4 Orbital Error Analysis Based on Relative Motion Theory	41
2.4.1 Initial Error Propagation Based on Algebraic Models	43
2.4.2 Correlation Characterization Based on Geometrical Model	65
References	74

3	Orbital Error Analysis Based on Historical Data	77
3.1	Overview of Error Analysis and Application Based on Historical Data.	78
3.1.1	Initial Covariance Analysis Based on Historical Data	79
3.1.2	Covariance Propagation Function Fitting Based on Historical Data	79
3.1.3	Promotion of Long-Term Precision Using Historical Data	81
3.2	Generation of Residual Data Considering Periodicity	83
3.2.1	Selection of Space Objects	83
3.2.2	Generation of Residual Data	83
3.3	Residual Data Preprocessing.	86
3.3.1	Residual Data's 2D Binning	86
3.3.2	Outlier Detection of Residual Data	88
3.4	Form of Fitting Function Considering Periodicity.	88
3.4.1	Error of Orbital Elements and Positions	88
3.4.2	Error Propagation Results of C-W Equations	89
3.4.3	Poisson Series	90
3.5	Error Coefficient Matrix Fitting	94
3.5.1	Least Squares Fitting of Poisson Series.	94
3.5.2	Determination of Poisson Series' Order	95
3.6	Results and Analysis	96
3.6.1	Fitting Result of Position Error	96
3.6.2	Effect of Poisson Coefficient Matrix	97
3.6.3	Error Analysis on U.S.-Russian Satellite Collision	101
3.7	Validation of Negative Correlation of Position and Velocity Error	102
	References.	104
4	Close Approach Analysis Between Space Object	107
4.1	Analytical Methods to Determine Close Approach	108
4.1.1	Hoots's Method.	108
4.1.2	Collision Detection Based on Secular Orbital Model	116
4.2	Numerical Methods to Determine Close Approach.	124
4.2.1	A-R Algorithm	125
4.2.2	A-N Algorithm	128
	References.	132
5	Calculation of Collision Probability	135
5.1	Calculation Method of Collision Probability and Simplification.	136
5.1.1	Overview of Calculation Method of Collision Probability.	136
5.1.2	Simplification of Calculation Method of Collision Probability.	141

- 5.2 Explicit Expression of Collision Probability in Circular Orbit 143
 - 5.2.1 Conjunction Geometry and Distance Analysis. 145
 - 5.2.2 Transformation and Projection of Error Covariance 148
 - 5.2.3 Explicit Expression of Collision Probability 151
 - 5.2.4 Examples 153
- 5.3 Explicit Expression of Collision Probability in General Orbit. . . . 157
 - 5.3.1 Explicit Expression of P_c in Terms of Conjunction Geometries 158
 - 5.3.2 Explicit Expression of P_c in Terms of NTW Components 167
 - 5.3.3 Examples 173
- 5.4 Explicit Expression's Eccentricity Bound 175
 - 5.4.1 Flight-Path Angle and Velocity's Magnitude of Eccentric Orbit 175
 - 5.4.2 Effects of Non-zero Flight-Path and Unequal Velocity 177
 - 5.4.3 General Effects of Eccentric Orbit. 178
- References. 181
- 6 Application of Collision Probability 185**
 - 6.1 Sensitivity Analysis of Collision Probability. 185
 - 6.1.1 Definition of Sensitivity of Collision Probability. 186
 - 6.1.2 Sensitivity Analysis Based on Explicit Expression 187
 - 6.1.3 Examples 193
 - 6.2 Maximum Collision Probability 201
 - 6.2.1 Introduction. 201
 - 6.2.2 Fixed Orientation 204
 - 6.2.3 Variable Orientation 213
 - 6.2.4 Examples 218
 - 6.2.5 Discussion. 224
 - 6.3 Missing- and False-Alarm Analysis of Conjunction Assessment. 226
 - 6.3.1 Essential of Conjunction Assessment and Misjudgment 227
 - 6.3.2 Safety- and Danger-Region of Conjunction Assessment 228
 - 6.3.3 Calculation of Probability of Missing Alarm. 231
 - 6.3.4 Calculation of Probability of False Alarm. 233
 - 6.4 Comprehensive Assessment of Collision Risk 235
 - 6.4.1 Definition of the F -Value 236
 - 6.4.2 Risk- and Quality-Assessment Parameters. 237
 - 6.4.3 Membership Functions of Assessment Parameters. 238
 - 6.4.4 Weighting Coefficients of Assessment Parameters. 241
 - 6.4.5 Examples 243
- References. 245

7 Orbital Anomaly and Space Events Analysis 247

7.1 Overview of Orbital Anomaly and Space Events 247

7.1.1 Orbital Anomaly and Space Events. 247

7.1.2 Orbital Anomaly Detection Based on Historical Data 249

7.1.3 Basic Procedure of Detecting Orbit Anomaly 252

7.2 Orbital Anomaly Detection Based on MWCF 254

7.2.1 Selection of Characteristic Orbit Elements 254

7.2.2 Generation of Dispersion Data 256

7.2.3 Detection of Outliers in Dispersion Data. 259

7.3 Orbital Anomaly Detection Based on Prediction Dispersion. 264

7.3.1 Secular Model of Characteristic Orbit Elements 265

7.3.2 Generation and Outlier Detection of Dispersion Data 266

7.3.3 Examples 268

7.4 GEO Orbit Anomaly Detection Based on Dispersion
of Drift Rate. 271

7.4.1 Characteristic Orbit Parameters and Variation Law. 272

7.4.2 Generation of Dispersion Data of Drift Rate. 277

7.4.3 Detection of Outliers in Dispersion Data. 281

7.4.4 Examples 283

7.5 Assessment of Detection Method of Orbital Anomaly 284

7.5.1 Discriminant Analysis and Misjudgment Probability. 287

7.5.2 Probability of False Alarm 288

7.5.3 Probability of Missing Alarm 293

7.5.4 Orbit Precision Requirement. 295

References. 296

Appendix A: Elements of T-H Equation State-Transition Matrix. 299

**Appendix B: Expression for Covariance Matrix Calculated
with C-W Equation.** 303

Appendix C: Some Figures of Chapter 3 305

Bibliography 317

List of Figures

Figure 1.1	Distribution of Space Objects	2
Figure 1.2	TLE Data Format Description	12
Figure 2.1	Satellite-based RSW and NTW coordinate systems	24
Figure 2.2	Schematic diagram of errors during space object orbital determination and prediction	26
Figure 2.3	Propagation of initial errors	26
Figure 2.4	Orbital model errors	27
Figure 2.5	Error propagation curve of object ①	38
Figure 2.6	Error propagation curve of object ②	39
Figure 2.7	Error propagation curve of object ③	40
Figure 2.8	Error propagation curve of object ④ (one day later)	41
Figure 2.9	Error propagation curve of object ④ (two days later)	42
Figure 2.10	Error propagation curve of object ③ in NTW coordinate system	43
Figure 2.11	Propagation curve of error standard deviations of object ①	53
Figure 2.12	Propagation curve of error standard deviations of object ②	53
Figure 2.13	Propagation curve of error standard deviations of object ③	54
Figure 2.14	Relative error curve of standard deviations of object ①	54
Figure 2.15	Relative error curve of standard deviations of object ②	55
Figure 2.16	Relative error curve of standard deviations of object ③	55
Figure 2.17	Propagation curve of error standard deviations of object ④ (in RSW coordination system)	59
Figure 2.18	Relative error curve of standard deviations of object ④ (in RSW coordination system)	59
Figure 2.19	Distribution of velocity errors using Monte-Carlo method at 1290s	60
Figure 2.20	Modifications of error data using orbit curvature	60

Figure 2.21 Propagation curve of error standard deviations of object ④ (in RSW coordination system with curvature modifications) 61

Figure 2.22 Relative error curve of standard deviations of object ④ (in RSW coordination system with curvature modifications) 62

Figure 2.23 Propagation curve of error standard deviations of object ④ (in NTW coordination system) 63

Figure 2.24 Propagation curve of error standard deviations of object ④ (in NTW coordination system) 63

Figure 2.25 Error ellipsoid of elliptical orbit 64

Figure 2.26 The changing curve of velocity angle Θ of object ④ with its passing the Perigee time τ 64

Figure 2.27 Correlation coefficients of LEO object based on C-W equations 71

Figure 2.28 Correlation coefficients of MEO object based on C-W equations 71

Figure 2.29 Correlation coefficients of GEO object based on C-W equations 71

Figure 2.30 Propagation of initial covariance (1 day) 73

Figure 2.31 Propagation of initial covariance (10 day) 73

Figure 3.1 The scheme for increasing long-term precision using historical data. 82

Figure 3.2 Generation of residuals between predicted and reference state in whole period 84

Figure 3.3 Generation of i th TLE's prediction residual to subsequent TLEs. 85

Figure 3.4 Residual data two-dimensional binning (**a left**, before binning; **b right**, after binning) 87

Figure 3.5 In-track (U) position error standard deviations for each data bins of four objects 87

Figure 3.6 Error standard deviation curves derived from Poisson series only consists of polynomial terms 98

Figure 3.7 Error standard deviation curve surface derived from Poisson series consists of both polynomial and trigonometric terms 99

Figure 3.8 Profile of in-track, normal, and cross-track error standard deviations curved surface of HEO object with respect to mean anomaly while $\Delta t = 5$ days 99

Figure 3.9 3σ error ellipsoid of Cosmos-2251 102

Figure 3.10 3σ error ellipsoid of Iridium-33. 102

Figure 3.11 Correlation coefficient curves of LEO object. 103

Figure 3.12 Correlation coefficient curves of MEO object 103

Figure 3.13 Correlation coefficient curves of GEO object 104

Figure 4.1 Relative distance between two objects 108

Figure 4.2 Geometrical prefilter 110

Figure 4.3 Relative orientation of orbital planes 111

Figure 4.4 Satellite positions in orbital planes 111

Figure 4.5 Angular windows 115

Figure 4.6 The value of the alert range ρ 117

Figure 4.7 Residual of the SGP4 model including only
the secular terms 118

Figure 4.8 Residual of the SDP4 model including only
the secular terms 119

Figure 4.9 Points of intersection of spherical projection
of two orbits. 120

Figure 4.10 Northern point of intersection of spherical projection
of two orbits. 121

Figure 4.11 The time differences of two objects crossing
the northern point and southern point 123

Figure 4.12 The orbital elements two objects crossing
the northern point and southern point 123

Figure 5.1 Description of the problem of P_c calculation. 136

Figure 5.2 Parameters of P_c density function and radius
of collision cross-section area on the conjunction plane 141

Figure 5.3 Conjunction always occurs in the vicinity of points
of closest approach between orbits 144

Figure 5.4 Circular orbit assumption and the conjunction
configuration 144

Figure 5.5 The closest approach vector between orbits $\Delta \mathbf{r}_{rad}$
and the closest approach vector at TCA $\Delta \mathbf{r}_{tca}$ 145

Figure 5.6 Definition of the encounter coordinate system. 146

Figure 5.7 Definition of the RSW coordinate system
of primary object 147

Figure 5.8 Closest approach distance in the horizontal plane 147

Figure 5.9 Position error covariance ellipsoids of two objects
which are aligned with the radial, in-track
and cross-track directions 148

Figure 5.10 Combined error variance in the horizontal plane σ_{sw}
versus the angle between two velocity vectors
when $\sigma_S > \sigma_W$ and $\sigma_S < \sigma_W$ 151

Figure 5.11 3σ combined error ellipsoid 155

Figure 5.12 The combined error ellipse and combined collision
cross-section on the conjunction plane 156

Figure 5.13 P_c versus error scaling factor using three methods. 157

Figure 5.14 Collision between space objects always occurs
in the vicinity of line of intersection of their orbit planes . . . 158

Figure 5.15 The necessary condition of closest approach between two orbits is that the relative position vector is perpendicular to both objects' velocity vectors. 159

Figure 5.16 The closest approach vector $\Delta \mathbf{r}_{\min}$ between orbits and the closest approach vector at TCA $\Delta \mathbf{r}_{\text{tca}}$ 160

Figure 5.17 Closest approach distance in the velocity plane. 163

Figure 5.18 Definition of the encounter coordinate system and position error ellipsoids of two objects. 164

Figure 5.19 The conjunction geometry and coordinate systems. 168

Figure 5.20 Transformation from modified NTW systems to the encounter system 171

Figure 5.21 P_c versus error scaling factor using three methods (ISS event). 175

Figure 5.22 Values for the flight-path angle θ and the ratio of velocity's magnitude η with respect to true anomaly when $e = 0.7$ 176

Figure 5.23 The P_c calculated by explicit expression and modified expression if the conjunction occurs at the perigee (a) and apogee (b) of secondary orbit 177

Figure 5.24 The P_c calculated by explicit expression and modified expression if the conjunction occurs at the position on secondary orbit where $\theta = \pm \sin^{-1} e$ 178

Figure 5.25 The differences between P_c calculated by explicit expression and modified expression with various true anomalies. 179

Figure 5.26 Cumulative distribution function of eccentricity of 12,043 objects in LEO ($h_a < 5000$ km) 180

Figure 5.27 The relative error between P_c calculated by explicit expression and modified expression with various true anomalies 180

Figure 6.1 Curve of combined errors in a horizontal plane varying with orbital plane included angles under two conditions. 191

Figure 6.2 Curve of P_c and its sensitivities changing with R-direction close approach distance 194

Figure 6.3 Curve of P_c and its sensitivities changing with S-direction close approach distance 194

Figure 6.4 Curve of P_c and its sensitivities changing with W-direction close approach distance. 195

Figure 6.5 Curve of P_c and its sensitivities changing with the combined error standard deviation in the direction R 196

Figure 6.6 Curve of P_c and its sensitivities changing with the combined error standard deviation in the direction S 196

Figure 6.7 Curve of P_c and its sensitivities changing with the combined error standard deviation in the direction W 197

Figure 6.8 Curve of P_c and its sensitivities changing with the conjunction angles 198

Figure 6.9 Curve of P_c and its sensitivities changing with the conjunction angles ($\sigma \times 5$). 199

Figure 6.10 Curve of P_c and its sensitivities changing with the conjunction angles ($\sigma \times 10$). 199

Figure 6.11 Curve of P_c and its sensitivities changing with combined radius r_A 200

Figure 6.12 The variation of ellipse's **a** size, **b** shape and **c** orientation in two-dimensions plane 201

Figure 6.13 Eight covariance situations for maximum collision probability assessment 204

Figure 6.14 Collision probability dilution assessment 207

Figure 6.15 Parameters of P_c density function and radius of collision cross-section area in the conjunction plane 209

Figure 6.16 Two-dimensional distribution (*left*) and three-dimensional mesh surface (*right*) of denary logarithm of P_{cmax} with respect to μ_x and μ_y when $r_A = 10$ m 211

Figure 6.17 The maximum and minimum P_c with respect to the angle θ 214

Figure 6.18 The limit case when $\sigma_y \rightarrow 0$ 216

Figure 6.19 The P_{cmax} and corresponding k indicate that the current state is not in the probability dilution region. 220

Figure 6.20 Two-dimensional P_c distribution with color determined by magnitude of P_c 221

Figure 6.21 Three-dimensional P_c mesh surface with respect to combined error standard deviations σ_R and σ_{SW} observed from different direction. 221

Figure 6.22 The curve of P_c with respect to the angle θ (Situation no. 5) 222

Figure 6.23 The curve of P_c with respect to combined error standard deviations σ_x (Situation no. 6) 223

Figure 6.24 The curve of P_c with respect to combined error standard deviations σ_x (Situation no. 8) 223

Figure 6.25 The P_c and P_{cmax} of five situations 225

Figure 6.26 Ellipses of constant collision probability 229

Figure 6.27 Safety- and danger-region of conjunction assessment. 230

Figure 6.28 Curve of size of boundary ellipse versus probability threshold P_T 230

Figure 6.29 Missing alarm region of collision assessment 231

Figure 6.30 Missing alarm region when the true value is in the origin 232

Figure 6.31 Curve surface of P_m versus (x_{true}, y_{true}) 233

Figure 6.32 False alarm region of collision assessment 234

Figure 6.33 Curve surface of P_{fa} versus (x_{true}, y_{true}) 235

Figure 6.34 Comprehensive assessment of collision risk based on the F -value 238

Figure 6.35 Membership function curve of miss distance. 239

Figure 6.36 Membership function curve of collision probability 240

Figure 6.37 Membership function curve of times to TCA 241

Figure 6.38 Membership function curve of quality of orbit determination 241

Figure 6.39 Evolution curve of RSW miss distance 244

Figure 6.40 Evolution curve of collision probability 244

Figure 6.41 Evolution curve of the f -value of quality assessment parameters 244

Figure 6.42 Evolution curve of the f -value of risk assessment parameters and the F -value. 245

Figure 7.1 Basic procedure of detecting orbit anomaly 253

Figure 7.2 Curve of semi-major axis, energy, angular momentum, and inclination versus time (Satellite Terra) 255

Figure 7.3 Moving window curve fitting 256

Figure 7.4 Flow chart of moving window curve fitting 258

Figure 7.5 Mahalanobis distance between the dispersion data and distribution center of satellite Terra 261

Figure 7.6 Mahalanobis distance between the dispersion data and distribution center of satellite ERS-2 262

Figure 7.7 Detection results of orbit outliers based on Mahalanobis distance (Terra) 262

Figure 7.8 Detection results of orbit outliers based on Mahalanobis distance (ERS-2) 263

Figure 7.9 Detection of Terra orbit anomalies in 2010. 263

Figure 7.10 Detection of ERS-2 orbit anomalies in 2010. 264

Figure 7.11 Schematic diagram of orbit prediction based on SGP4 secular term model 267

Figure 7.12 Jitter of characteristic orbit parameters 267

Figure 7.13 Detection flow of orbital anomalies based on prediction dispersion. 268

Figure 7.14 Variations and maneuvers of Terra’s semi-major axis and inclination in 2010 269

Figure 7.15 Smoothing of cataloged and predicted values of Terra’s semi-major axis in 2010. 270

Figure 7.16 Detection results of Terra maneuvering in 2010 270

Figure 7.17 Dispersion distribution of Terra’s semi-major axis and inclination in 2010 271

Figure 7.18 Definition and components of inclination vector and eccentricity vector 273

Figure 7.19 Theoretical drift rates of mean longitude and inclination of Beidou-1B satellite 277

Figure 7.20 Theoretical drift rates of mean longitude and inclination of STTW-1 satellite 278

Figure 7.21 Initial and smoothed data of semi-major axis of Beidou-1B satellite 279

Figure 7.22 Theoretical drift rate of mean longitude of Beidou-1B satellite calculated based on initial and smoothed data 280

Figure 7.23 Detection of outliers in drift rate dispersion based on Mahalanobis distance (Beidou-1B) 280

Figure 7.24 Detection of outliers in drift rate dispersion based on Mahalanobis distance (STTW-1). 281

Figure 7.25 Distribution of drift rate dispersions at uncontrolled sections of STTW-1 satellite. 282

Figure 7.26 Detection results of orbital maneuvers of Beidou-1B satellite 284

Figure 7.27 Detection results of orbital maneuvers of STTW-1 satellite. 285

Figure 7.28 Flow chart of GEO orbit anomaly detection based on dispersion of drift rate. 286

Figure 7.29 Curve of Pn and Pfa changing with Mahalanobis distance threshold 290

Figure 7.30 d_M Probability density function in terms of different n values 292

Figure 7.31 d_M Probability distribution function in terms of different n values 292

Figure 7.32 Probabilities of false and missing alarm 294

Figure 7.33 Probabilities of false and missing alarm (one-dimensional case). 294

Figure 7.34 Curve of missing alarm probability changing with orbital errors 295

Figure C.1 TLE’s prediction residual state components in UNW coordinate system versus temporal difference for LEO object. 305

Figure C.2 TLE’s prediction residual state components in UNW coordinate system versus mean anomaly for LEO object 306

Figure C.3 TLE’s prediction residual state components in UNW coordinate system versus temporal difference for MEO object 306

Figure C.4 TLE’s prediction residual state components in UNW coordinate system versus mean anomaly for MEO object 307

Figure C.5 TLE’s prediction residual state components in UNW coordinate system versus temporal difference for HEO object 307

Figure C.6 TLE’s prediction residual state components in UNW coordinate system versus mean anomaly for HEO object 308

Figure C.7 TLE’s prediction residual state components in UNW coordinate system versus temporal difference for GEO object 308

Figure C.8 TLE’s prediction residual state components in UNW coordinate system versus mean anomaly for GEO object 309

Figure C.9 In-track (U), normal (N), and cross-track (W) error standard deviation’s Poisson series fitting curved surfaces of LEO objects ($n = 2, m = 4$). 310

Figure C.10 In-track (U), normal (N), and cross-track (W) error standard deviation’s Poisson series fitting curved surfaces of MEO objects ($n = 2, m = 4$) 311

Figure C.11 In-track (U), normal (N), and cross-track (W) error standard deviation’s Poisson series fitting curved surfaces of HEO objects ($n = 2, m = 4$). 312

Figure C.12 In-track (U), normal (N), and cross-track (W) error standard deviation’s Poisson series fitting curved surfaces of GEO objects ($n = 2, m = 4$). 313

Figure C.13 TLE’s prediction residual state components in RSW coordinate system versus temporal difference for LEO object (20 days). 314

Figure C.14 TLE’s prediction residual state components in RSW coordinate system versus temporal difference for MEO object (20 days) 314

Figure C.15 TLE’s prediction residual state components in RSW coordinate system versus temporal difference for GEO object (20 days) 315

List of Tables

Table 2.1	Types of objects and orbits	37
Table 2.2	Orbital elements of objects.	38
Table 2.3	Types of objects and orbits	51
Table 2.4	Orbital elements of objects.	51
Table 2.5	Position and velocity coordinates in ECI coordinate system.	52
Table 2.6	Epoch and initial orbital elements	72
Table 2.7	Initial position and velocity covariance matrix	72
Table 3.1	List of space objects selected for the study.	83
Table 3.2	Poisson coefficient matrix of in-track (U) error standard deviation for four objects.	97
Table 3.3	The temporal and angular parameters of U.S.-Russian satellite collision.	101
Table 3.4	Position error standard deviations of two satellites	101
Table 4.1	Orbital types and parameters of objects	117
Table 4.2	TLEs of two space objects.	123
Table 5.1	Position and velocity coordinates in the ECI system and eccentricities	154
Table 5.2	Conjunction geometry of the U.S. and Russian satellite collision event	155
Table 5.3	Positional error standard deviations in radial, in-track, cross-track direction and the combined error.	155
Table 5.4	Collision probabilities from Chan's method and explicit expressions.	156
Table 5.5	Position and velocity coordinates in the ECI system and eccentricities (ISS event)	173
Table 5.6	Conjunction geometry of the ISS event	174
Table 5.7	Positional error standard deviations in normal, tangential, cross-track direction and the combined error (ISS event)	174
Table 5.8	Collision probabilities from Chan's method and explicit expressions (ISS event)	174

Table 5.9 Relative error for two cases in various eccentric orbits 180

Table 6.1 Two types of sensitivities to R , S and W 188

Table 6.2 Two types of sensitivities to standard deviations of errors in directions R , S and W 189

Table 6.3 Two types of sensitivities to conjunction angle. 190

Table 6.4 Determination of the sign of sensitivity to conjunction angles 192

Table 6.5 Expressions of sensitivities of P_c to various influencing factors. 193

Table 6.6 P_c sensitivities value to components R , S and W 195

Table 6.7 P_c sensitivity value to combined error standard deviations in directions R , S and W 197

Table 6.8 P_c sensitivity value to the conjunction angle φ 199

Table 6.9 P_c sensitivity value to combined radius r_A 200

Table 6.10 The P_{cmax} calculated using different relative position. 211

Table 6.11 The maximum and minimum P_c with respect to the angle θ 214

Table 6.12 Summary of the results 219

Table 6.13 P_{cmax} and corresponding error scaling factor and standard deviations (Situation no. 2) 220

Table 6.14 The P_{cmax} and corresponding combined error standard deviations (Situation no. 4) 220

Table 6.15 P_{cmax} and corresponding combined error standard deviations in arbitrary error ellipse shape case computed using empirical mesh surface, explicit expression, and approximate expression 221

Table 6.16 The P_{cmax} and corresponding combined error standard deviations (Situation no. 6) 222

Table 6.17 The P_{cmax} and corresponding combined error standard deviations (Situation no. 8) 223

Table 6.18 Summary of the resulted P_{cmax} and corresponding error covariance of five situations 224

Table 6.19 Judgment matrix of risk assessment parameters. 243

Table 6.20 Judgment matrix of quality assessment parameters 243

Table 6.21 Weighting coefficient of each parameter. 243

Table 6.22 Risk assessment parameters and quality assessment parameters of US-Russian satellite collision event. 243

Table 7.1 Orbital maneuvers of satellite Terra In 2010. 269

Table 7.2 Orbital parameters of space objects 272

Chapter 1

Introduction

The notion of space object generally refers to all artificial objects operating and traveling in outer space, including man-made earth satellite, space station, space shuttle, and space debris produced from them or their wastes. A majority of space objects are space debris in quantitative terms. In accordance with the definition concluded at the 32nd Conference of UNCOPUOS, STSC, the space debris (alias: the orbital debris) are the artificial objects and its fragments, debris, and components and parts, which are out of service, cannot be retained or cannot have the original functions resumed, traveling along earth orbits or reentering the condensed atmospheric layer.

Along with mankind's continuous strides made in aerospace industry and frequent and intensive launching of space activities, the quantity of in-orbit space objects (spacecrafts, rocket bodies, and space debris) is increasing, showing the sign of aggravation of space debris environment. As in January 2014, the total number of trackable space objects catalogued in U.S. Space Surveillance Network (SSN) has reached 16655, among which 3715 are with payload (including the operating ones and disabled ones), 12940 are rocket bodies and space debris, taking up 3/4 of the total and such number is still on the increase [1]. It is estimated that more than 300,000 space objects with a diameter of greater than 1 cm are doing orbital travel and millions more with a diameter smaller than or equal to 1 cm are still out in the space. See Fig. 1.1 for the distribution of the bigger space objects with definitive orbits.

The increase of space objects markedly enlarges the possibility of collisions between spacecraft and space objects, putting the normal operations of spacecraft, sustainable development of aerospace industry and terrestrial space environment into great jeopardy with a long-lasting and extensive threat. There are several space objects collisions in history. For instance, ESA Ariane rocket debris ran upon the gravity gradient stabilizing arm [2, 3] of French Cerise satellite on July 24th 1996; the last stage of U.S. Thor Burner 2A rocket collided with the last stage debris of Chinese CZ-4 carrier rocket in the sky 885 km to the Antarctic [4] on January 17th 2005. All these collisions were taken place between debris or between debris and

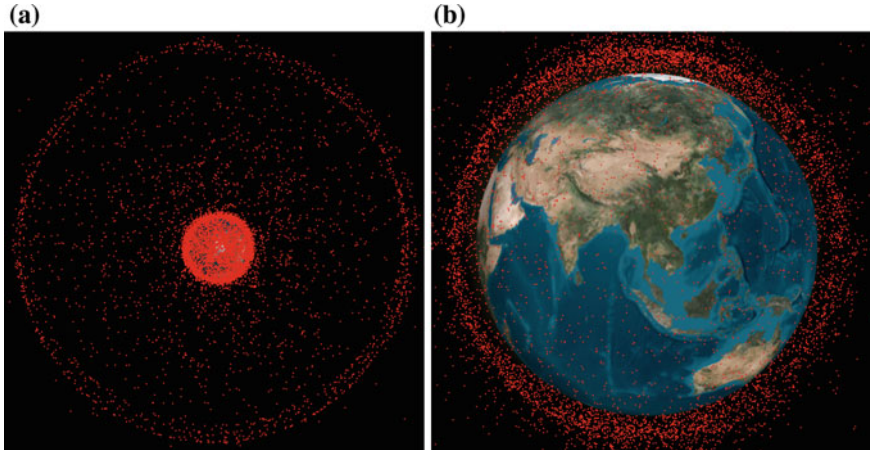


Fig. 1.1 Distribution of Space Objects

satellites. On February 10th 2009, for the first time in human history, the collision between two unbroken in-orbit satellites (the U.S. Iridium-33 satellite (SSN NO. 24946, international designator 1997-051C) and Russian disabled communication satellite Cosmos-2251 (SSN NO. 22675, international designator 1993-036A) was happened in the sky 790 km to Siberia [5]. Plenty of space debris pieces were produced during collision. Due to higher altitude of collision point, the debris will be in orbit for a long time, constantly deteriorating the space environment. Thus, the in-orbit collision between the U.S. and Russian satellites has sounded an alarm bell to mankind for the space safety.

The U.S. National Space Policy Intersector Guideline [6] includes 7 articles, as prescribed in Article 3 of “Preserving the Space Environment and the Responsible Use of Space”:

Preserve the Space Environment. For the purposes of minimizing debris and preserving the space environment for the responsible, peaceful, and safe use of all users, the United States shall:

- Lead the continued development and adoption of international and industry standards and policies to minimize debris, such as the United Nations Space Debris Mitigation Guidelines;
- Develop, maintain, and use space situational awareness (SSA) information from commercial, civil, and national security sources to detect, identify, and attribute actions in space that are contrary to responsible use and the long-term sustainability of the space environment;
- Continue to follow the United States Government Orbital Debris Mitigation Standard Practices, consistent with mission requirements and cost effectiveness, in the procurement and operation of spacecraft, launch services, and the conduct of tests and experiments in space;

- Pursue research and development of technologies and techniques, through the Administrator of the National Aeronautics and Space Administration (NASA) and the Secretary of Defense, to mitigate and remove on-orbit debris, reduce hazards, and increase understanding of the current and future debris environment; and
- Require the head of the sponsoring department or agency to approve exceptions to the United States Government Orbital Debris Mitigation Standard Practices and notify the Secretary of State.

Foster the Development of Space Collision Warning Measures. The Secretary of Defense, in consultation with the Director of National Intelligence, the Administrator of NASA, and other departments and agencies, may collaborate with industry and foreign nations to: maintain and improve space object databases; pursue common international data standards and data integrity measures; and provide services and disseminate orbital tracking information to commercial and international entities, including predictions of space object conjunction.

What specified above need to be aided with Space Situational Awareness which is able to acquire space objects and space environment information through continuous observation of the space and information gathering to strengthen in-orbit safety of spacecrafts, space event analysis and assessment, and threat forecast and assessment. Satellite maneuver, anticipated and unanticipated launch and reentry and task-hindered space environmental change are mainly described in the space events.

It is imperative that only by grounding on real-time and valid space object data can the space situation be more explorable and controllable through Space Situational Awareness. But how to acquire space object data, what kind of space object data can be obtained, and how to use such data are the issues to be further focused on in Space Situational Awareness studies.

1.1 Data Source—Space Surveillance System

The responsibility of discovering and detecting all space objects and their orbits are undertaken in Space Surveillance as a key component of Space Situational Awareness. By various observational means, motion parameters of space object in real time are able to be discovered, tracked and measured by Space Surveillance network, so that its orbital characteristics can be determined. Cataloging large space objects and updating their orbital data from time to time according to the observed and determined orbital data, the dynamic database of space objects can be created.

1.1.1 Space Surveillance

The tasks of Space Surveillance are to detect, track, identify, and catalogue all in-orbit space objects as well as to provide target catalog database and real-time orbital and precise characteristics information. The space object indicated in this book refers to the space object which has its orbit determined by cataloguing. Space Surveillance is one of main measures to obtain the space situational information; the Capacity of Space Surveillance is the rock upon which the space superiority in future is built and the key basis for the space control. Moreover, the Space Surveillance is also taken as an important aerospace infrastructure and basic capability, playing significant roles in both military and civil sectors.

Space Surveillance is consisted of 4 tasks categorized as below [7]:

- (1) To maintain the space object catalog and detect new launches in a timely manner and to monitor orbital maneuvers, deorbit and disintegration of existing objects;
- (2) To identify space objects, evaluate task load and analyze threat level;
- (3) To initiate space control support, oversee the performance of space control related terms and conditions and provide satellite attack warning;
- (4) To provide key spacecrafts with collision-proof support.

Among those four tasks listed above, the Space Object Cataloguing is the basic task which must be completed with the Space Surveillance System. The Space Surveillance system shall be capable of acquiring space object orbital (trajectory) data and providing necessary measured data for orbital determination so as to present the orbital parameters to reflect target's motion track and serve the Space Object Cataloguing.

The Space Object Catalog Maintenance is with three implications [7]:

- (1) Update catalog target orbital element;
- (2) Prompt detection of new launch;
- (3) Prompt detection of existing target's orbital variation, i.e., orbital transfer, deorbit or disintegration.

Orbital observation, orbital determination, orbital prediction, and error analysis are included in the orbit-related tasks in Space Surveillance. These tasks shall be completed based on a space object dynamic catalog database and a complete set of forecasting methods which shall be deemed as a complete set integrated with hardware system and software system covering observation data acquisition strategy, observation data processing method, orbital determination method, orbit prediction method, target history, and current database.

1.1.2 Space Surveillance System

The Space Surveillance system is a national strategic information acquisition system [7] applied to monitor and observe the processes of artificial celestial body entering, operating and existing in the space in order to acquire the orbital, functional and status information of such artificial celestial body. The mankind has undergone the phases of traveling into the space, space utilization, and space observation and is striding forward to an age when people have control over the outer space with the purpose of preventing hostile forces from entering and taking advantages of the outer space while ensuring its own forces and alliances being capable of entering and using the outer space. As one of indispensable foundations for space control, the Space Surveillance System is deemed as a strategic force of a nation.

The space surveillance network, space surveillance center, and timing and communication auxiliary systems are generally assembled in the Space Surveillance system. The space surveillance network is the basis for space surveillance with data acquisition function. The space surveillance center, the core part for space surveillance, undertakes the missions like tasks planning, data processing for orbital determination, identification, cataloguing, assessment and database building.

The process of tracking and orbital determination of a space object is a procedure of estimation to the status of the space object or its orbital elements based on the measured data with errors obtained from ground-based and space-based observatories and imprecise astrodynamics equations. The space object orbital estimation was divided into two conditions: one is for one's own satellite by means of collaborative working mechanism, generally known as Spacecraft Observation and Control; the other is the acquisition of orbital, status, and functional information of non-collaborative satellite and space debris by means of monitoring or space surveillance.

As required in Space Surveillance tasks, data from all objects within the designed scope should be independently, reliably, and periodically collectible with space surveillance network, so that all objects within the scope can be cataloged and updated to discover new launches and orbital variations of existing objects. The independent, reliable, and periodical data acquisition of the space surveillance network is known as the essential requirement.

In order to track and measure all objects meeting the criterion scope, the measuring equipments for space surveillance should be operated based on non-collaborative working mechanism. For now, the Space Surveillance equipments based on non-collaborative working mechanism include radar apparatuses and optical devices, which operate based on active reflective tracking and passive receiving in respective. Due to the different types and features of these two kinds of equipment, they play different roles in the space surveillance missions.

The radar apparatuses mainly refer to large-size mechanical radar or phased array radar. The advantages of radar apparatuses:

- (1) Wide range of continuous search: single phased array is able to cover a search area of $120 \times 80^\circ$ and above;
- (2) Multi-target real-time tracking: being able to track more than 100 objects simultaneously;
- (3) All-weather service: being able to operate at all times in any illumination condition or under any weather condition;
- (4) Completely measured data: multiple types of data such as distance measurement, angle measurement, speed measurement, and echo signal;
- (5) High measurement accuracy: general distance measurement is with accuracy of 30 m, general angle measurement is with accuracy of 1 mrad and speed measurement is with accuracy of 0.1 m/s.
- (6) Good real-time performance: direct output of tracking data for rapid processing.
- (7) Strong orbital determination ability: orbital determination with precision of kilometer magnitude can be obtained based on short-arc data from a single station.

The disadvantages of radar apparatuses:

- (1) Limited operating range: being difficult to track target from a distance of over 5000 km.
- (2) High cost: a single set of equipments with an operating range of over 3000 km costs no less than RMB 200 million.
- (3) Limited continuous working hours: due to higher transmission power, the continuous working hours of the equipment will be shortened if the heat generated from the transmitter fails to be emitted.
- (4) Weak imaging capability: space surveillance radar is unable to provide a quality-proven imaging service at present.

The optical devices mainly refer to the electro-optical/electro-optical telescope which is endowed with following advantages:

- (1) High angular accuracy: the pointing accuracy can reach $5''$ – $10''$ which is 1–2 orders of magnitude higher than the angular accuracy by radar;
- (2) Wide operating range: small-size optoelectronic device is able to track target with luminance of stellar magnitude over 11, while large-size optoelectronic device is able to track target with luminance of stellar magnitude 17 above, thus, it is applicable to deep-space tracking;
- (3) Strong imaging ability: the angular resolution can reach $0.1''$ – $0.2''$;
- (4) Lower cost, easy operation and maintenance and long service life.

The disadvantages of optical devices:

- (1) The device may be affected by illumination environment and weathers; its actual working hour is limited;
- (2) Small field range and guidance is required;
- (3) Only angular measurement ability is available.

As radar apparatuses and optional devices have both merits and drawbacks, both types of equipment are used and their advantages are integrated when constructing the Space Surveillance network, for instance:

- (1) radar apparatuses are mainly applied for regular surveillance on low-orbit space objects with supplementary functions of optical devices;
- (2) optical devices are mainly applied for surveillance on medium and high-orbit space objects;
- (3) optical devices are mainly applied for imaging identification.

The platforms for Space Surveillance system are generally ground-based and space-based. The ground-based surveillance network is consisted of ground-based radar and ground-based electro-optical telescope. The proven technologies, lower cost of investment and unrestricted size and mass limitations, big-caliber space antenna with higher transmission power can be adopted to search and track space objects more efficiently. However, the ground-based surveillance is subject to weather conditions and lack of maneuvering capability. Besides, The visual range and time are limited by earth curvature, so global station placement is required [8, 9]. It is a challenge to set up station beyond the borders.

The space-based Space Surveillance can be construed as optical observation or other forms of surveillance on space objects by taking advantage of space-based platforms. The space-based observation platform operates in orbit which is free from the atmospheric restriction. Geosynchronous earth orbit (GEO) satellite on the other side of the earth can be observed. The “optical visibility period” effect can also be minimized to increase the chance for observation and prolong the observation time. Its observation sensitivity can be enhanced by defying the earth’s atmospheric effects and under the circumstance of unavailability of earth’s atmosphere, the close-range and high-resolution imaging of space object can be achieved to obtain the image feature in order to further strengthen the space object observational and identification capacities. In spite of this, the space-based Space Surveillance system is with the disadvantages like high cost and complicated technologies, and its onboard processing (OBP) capability and power are also far less comparable than the ground-based surveillance system.

1.1.3 Space Surveillance Networks

The United States, Russia, and Europe have established Space Surveillance systems characterized by a broad observational area and scope. They conduct both periodical and continuous observations.

The U.S. has established efficient ground-based Space Surveillance system and is undergoing the development of space threat warning technology. The U.S. space surveillance system is evolved progressively from the end of the 1950s. So far, the Space Surveillance system including ground-based electro-optical devices and radar

devices based in homeland and covers the world has been established. The system consists of over 30 sets of radio sensors, optical sensors, space-based sensors of more than 20 types distributed in 16 different locations around the world, and 2 control centers (1 master control and 1 backup). Functionally speaking, both space surveillance and missile early-warning are put into account in the course of the system development by combination of ground and space distributions, and the observation stations have been deployed all over the world.

The sensors applied to the U.S. Space Surveillance nets can be subdivided into three types: exclusive sensors, compatible sensors and available sensors. Majority of the sensors are owned by the Air Force Space Command (AFSPC) and Army Space Command subordinated to the AFSPC, and non-government institutions only own minority of them. The mission division, coordination, command, and control of space surveillance tasks are performed and completed at the Cheyenne Mountain Space Control Center co-owned by North American Aerospace Defense Command (NORAD) and AFSPC. Information obtained by sensors is sent to the center for fusion, processing, cataloguing, and distribution. At present, space objects sized 10 cm on low earth orbits and space objects sized 30 cm on geosynchronous orbits can be observed respectively and over 16,000 man-made space objects have been tracked and catalogued. The center is now with the ability to identify majority of satellites in orbit.

The space-based visible (SBV) sensor onboard the U.S. Midcourse Space Experiment (MSX) satellite has had the space-based space surveillance technologies and functions verified. The SBV is able to cover the entire geosynchronous belt with the primary mission of collecting orbital measurements and luminosity information from all kinds of space objects. It was launched in April 1996, and after that, the previous 18 months were mainly spent in affirmation of space-based space surveillance concept and assessment of SBV sensor's performance. In October 1997, SBV sensor started to operate as a critical sensor instead of an experimental one in the space surveillance network. It officially became a constituent part of Space Surveillance network of the United States in October 2000, having the key object revisit rate increased by 50%.

Russia has also established a "Cosmic Space Surveillance System", including a Cosmic Space Surveillance center, for observing and tracking the space environment. However, the equipments adopted by Russia for Space Surveillance are mainly the electro-optical telescopes. The "sky window" system, a typical active ground electro-optical space surveillance and tracking system owned by the Russian space force, is located in the mountainous area at the borders of Tajikistan.

The Space Surveillance network which are devised in Europe comprise a GRAVES electromagnetic fence system for low-orbit space object observation and measurement and 3 telescopes for observation of space debris in geosynchronous orbits. The French Aeronautics & Astronautics Research Institute has been engaged in the research and development of GRAVES space surveillance radar since 1992.

The radar transmitter is installed in Dijon and the receiver is located in Provence. It was completed and performed the experimental detection for the first time at the end of 2004; the product in the first phase was officially delivered and put into service in November 2005. With the aid of this electromagnetic fence system, satellites at a height of lower than 1000 km can be detected and the time from satellite launch to being detected will not exceed 24 h, meeting the requirements of rapid detections of newly identified satellites, orbital maneuvers, and space debris produced from disintegration.

1.2 Data Type-Space Object Catalog

The space catalog data is the foremost product obtained through Space Surveillance. The essential requirements of space catalog data are:

- (1) Integrity: all targets above certain size shall be included;
- (2) Real-time: rapid cataloging of new targets and quick updating of orbital elements;
- (3) Precision: accurate information and high precision of orbital data;
- (4) Open-source: convenient and practical with guaranteed value of the data.

1.2.1 Basic Information of Space Catalog Data

The space catalog data shall at least include the basic information, orbital information and characteristics information of a space object.

- (1) Basic information

The basic information refers to the information which will remain unchanged or change very slowly when the object is launched into the space, including but not limited to:

- (a) Object no
- (b) Object name
- (c) Object type
- (d) Country
- (e) Launch date
- (f) Launching site
- (g) Deorbit date
- (h) Basic orbital information (period, inclination, perigee and apogee)

(2) Orbital information

The orbital information refers to the orbital elements, error information and orbital model information that describe the space object's orbital motion, including but not limited to:

(1) Orbital elements (or position and velocity vectors) which are subject to update from time to time:

- ① Mean elements
- ② Instantaneous elements

(2) Orbital covariance data:

- ① Rectangular coordinates
- ② Orbital elements

(3) Orbit and covariance forecasting model

- ① Analytical model
- ② Numerical model

(3) Characteristic information

In addition to the payload, operating status and external shape of the space object, the characteristic information also includes but not limited to:

- (1) Payload
- (2) State of service
- (3) Appearance structure
- (4) Radar Cross Section, RCS
- (5) Optical characteristic
- (6) Surface-to-mass ratio
- (7) Orbit controllability
- (8) Attitude controllability

1.2.2 Orbital Model of Catalog Data

Orbit prediction method is an important component of the Space Surveillance system. The primary tasks are to establish accurate and efficient orbit prediction model, promptly provide the long-term accurate orbit prediction data of space objects to determine the positions of the space objects within designated flight duration.

Both analytical and numerical orbit prediction methods are available in response to different requirements of application. Analytic method is to use analytic

functions to express the variations of orbital elements versus time and solve differential equations based on the initial orbital elements to obtain the analytical results. The merit of analytic method is that the change principles of elements versus time can be explicitly shown through the solution functions, and it is conducive to understand the effects of perturbations to the orbit elements along with high computational efficiency. The demerit of such method is the complicated derivational process and difficulties in derivation of high-order solution with lower precision than numerical solution. The analytic method commonly employed in China is the quasi-mean element method proposed by Liu [10], and that commonly used overseas is the SGP4/SDP4 model applied by the U.S. SSN [11–13].

The numerical method is a classical approach for orbital determination, namely, the integral solution of orbital dynamic model with numerical integration algorithm. Along with the improvement of accuracies of attitude and orbital determinations of the spacecraft, the orbital perturbation model is being unceasingly ameliorated, which leads to the continuous improvement of the precision of numerical solution. But the demerit of such method is that the inner variation principles of orbital elements versus time can hardly be obtained and long-term forecast needs to be calculated step by step. In China, Li [14], has made in-depth studies on numerical solution models and methods while World's leading achievements include the GEODYN model employed by Goddard Space Flight Center, NASA and High Precision Orbit Propagator (HPOP) applied to STK software [15, 16].

Moreover, Wu [17] has made exploration of the particularities of space debris orbital determination. As there exist a large amount of space debris known as passive targets with large and unknown surface-to-mass ratios, can hardly be observed precisely, the orbital determination of space debris usually depends on simple theories and approaches with more emphases on initial orbit determination and data association. The orbits of space debris are mainly applied for conjunction assessments, risk object deorbit forecasts, and long-term orbital evolutions of small debris, attaching importance to the forecast accuracy. However, the atmospheric model must be determined before orbit forecasting. Studies on orbital determination of space debris and atmospheric model must be carried on at the same time.

Moreover, the semi-analytic method which combines analytic method and numerical method [15, 16] together so far has not been well developed and has seldom been applied in engineering projects.

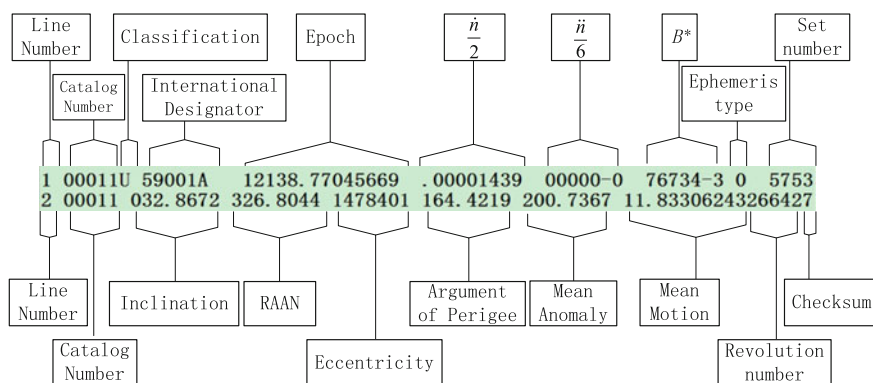
1.2.3 U.S. Catalog Data

The U.S. SSN is the most advanced space surveillance system in the world nowadays, with regular missions of cataloging in-orbit objects, maintaining catalog data, tracking in-orbit objects, and updating orbit elements. NORAD releases the Two Line Element (TLE) data of space objects at www.space-track.org with regular period intervals.

A TLE set is a data format encoding a list of mean orbital elements. The TLE data representation is specific to SGP4/SDP4 model, which is an analytic model. In December 1980, SSN published the literature [11], unfolding 5 sets of orbital models for forecasting, namely, the SGP, SGP4, SDP4, SGP8, and SDP8 and corresponding program source codes. In literature [11], all space objects are divided into near-earth objects (period of less than 225 min) and deep-space objects (period of greater than or equal to 225 min), orbit prediction models are also categorized into near-earth models and deep-space models accordingly. Among these 5 orbital models, SGP and SGP4 are for near-earth orbital determination with discrepancies in expression forms of mean angular velocity and drag; SDP4, extended from SGP4, is for deep-space orbital determination; SGP8 is also applied to near-earth orbital determination with different method of differential equation quadrature adopted; SDP8, extended from SGP8, is applied to deep-space orbital determination with deep-space influential model equation same as the one in SDP4. The SGP model is too simple and inaccurate. SGP/SDP4 is the mainstream in application at present, and the TLE data released through SSN are prepared for SGP4/SDP4 models.

TLE format uses two lines of 80-column ASCII text to store the orbital elements of space object and other relevant information. The specific format can be seen in Fig. 1.2.

It contains object-related information such as object number, international designator and atmospheric drag, 6 independent orbital elements ($i, \Omega, e, \omega, M,$ and n). The epoch defines the time reference point. In TLE format, B^* is an atmospheric drag-related item which describes the impact degree of the space object caused by the atmospheric drag. To obtain a forecasting result with high precision, the forecast model must have these perturbations restructured in the same way, thus, TLE data must be employed together with the SGP4/SDP4 orbit prediction model to forecast the status of space object. If TLE data are put into other models, the result with



Classification: U—publicly available, C—classified
 Ephemeris type: 0 for SGP4/SDP4 model

Fig. 1.2 TLE Data Format Description

lower precision will be obtained even though such model is of highly precision or is a numerical integration model.

Characteristics of TLE orbital catalog data are:

- (1) Integrity: the most comprehensive space object cataloging system in the current world;
- (2) Real-time: common target will be updated every 1–2 days, critical target will be updated 2–3 times every day;
- (3) Precision: orbital precision meets general analysis requirements;
- (4) Open-source data: data are open to all people and organizations except for the data of key military satellites of the United States and its alliances.

1.3 Data Application—Conjunction Assessment and Situational Analysis

Space Surveillance system produces a large amount of space object catalog data every day. Using these data, current spatial positions and velocities of space objects and those within a period of time in future can be calculated, which not only lay a foundation for the understanding of space situational distribution, but also provide strong support to in-orbit spacecraft collision warning and avoidance maneuvers.

The historical change process of a space object's orbit can be defined from catalog data accumulated in a long term. Data analysis can reveal principles behind, such as orbital maneuver events and debris flow characteristics. Innovative researches on such principles will facilitate space situational analysis and improve the understanding of the space object's in-orbit characteristics.

1.3.1 Conjunction Assessment

The space object conjunction assessment has drawn increasingly wide attentions in the international arena. In recent years, the rapid development of aerospace industry in China has made space object conjunction assessment increasingly important, particularly for the manned spacecraft and space station which operate on near-earth orbits characterized by dense concentration of space objects. The size of manned spacecraft or space station is larger than general satellites, and hence the risk of collision with space object is much higher. Space object conjunction assessment should be further studied to ensure the safety of manned spacecraft, and maintain the space safety. In spaceflight mission analysis and spacecraft long-term management, prevention from collision with space object has become a problem demanding prompt solution. A set of effective theoretical methods and operation strategies are presented to deal with such increasingly serious threats.

Spacecraft must perform collision avoidance maneuver to ensure safety if encountering space objects with large sizes (for example, larger than 10 cm). Such kinds of space objects could be tracked and catalogued for approaching analysis and conjunction assessment.

Space object conjunction assessment and collision avoidance maneuver are two relatively independent aspects in one complete process. The collision risk parameters (i.e. minimum distance, collision probability, and maximum collision probability) can be obtained by conjunction assessment, which is based on the orbital status and covariance error information in future. Designated norms are used to determine whether the space object is in dangerous zone (i.e. the approaching distance is less than the range threshold, or the collision probability is greater than the probability threshold). If it is, conjunction alarm will be given and corresponding measures shall be taken. Therefore, the conjunction assessment is a discriminant analysis based on risk assessment. Parameters obtained through analysis will tell us whether the spacecraft is exposed to risk or not, and collision avoidance maneuver need to be performed or not.

Considering the restrictions of mission and combining with orbit maintenance, the spacecraft is able to avoid the collision with space object by means of avoidance maneuvers in accordance with information, such as the collision risk parameters and approaching geometric relationship. Guidelines of collision avoidance maneuver performance and judgment standard are given based on studies of optimal collision avoidance maneuver method and implementation strategy. Hence, the collision avoidance maneuver, is simply an orbital transfer mission with core ideas of optimal orbital transfer and task restriction.

Due to the inevitable existence of errors in observation, orbital determination, and orbital prediction, the error analysis thus becomes an indispensable component of a space object collision assessment. In view of all error factors, the assessment criteria, such as minimum approaching distance and collision probability, are extensively considered in the space object collision assessment in response to the Box method and P_c method for conjunction assessment. According to the Box method, any space object that enters the pre-set and fixed evasion areas is considered as risk object, and decision for collision avoidance maneuver will be made in response. This method is an averaging method. The zoning of areas is based on the average error from space object orbital prediction which is excessively conservative in most cases and may result in useless collision avoidance maneuver operation and loss of maneuver capability. The P_c method is a collision risk assessment method with higher level of precision than the Box method. The collision probability not only depends on the minimum distance between the risk target and the spacecraft at the Time of Closest Approach (TCA), but also involves the position-velocity geometrical relationship and its uncertainties at the approaching moment of the two objects. Unlike the average error-based early-warning zoning of the Box method, the P_c method is developed to concrete error information with advantages of giving less collision avoidance maneuver commands and lower false alarm rate than the Box method. The collision probability-based space object

conjunction assessment is the mainstream analytical method for early warning applied in the current international community.

The collision probability-based space object conjunction assessment is developed under the following key techniques: space object dynamic orbital database; space objects approaching analysis; orbit prediction method and error analysis; collision probability calculating method (linear relative motion and non-linear relative motion); collision risk comprehensive assessment method and collision probability-based collision avoidance maneuvering method, etc.

Two theoretical issues are involved in terms of conjunction assessment on traceable space object based on Space Surveillance: one of them is about space object orbital theory, including orbital determination, orbital prediction, approaching analysis, evasion maneuver, etc.; the other is related to error and probability, such as the analysis on the inevitable uncertainties and probability problem aroused in orbital determination and orbit prediction as well as assessments on collision risk and evasion maneuver with the probability. Collision assessments rely on approaching distance and collision probability calculations, which are based on the parameters, such as the position-velocity vectors at the approaching moment of both objects, corresponding error covariance matrix, and their sizes and shapes. The forecast accuracy of these parameters will greatly affect the accuracy degree of collision probability and validity of conjunction assessment.

To deal with the increasing collision threats, applicable conjunction assessment and evasion maneuvering systems in engineering practices are started to be built or have already been established by some countries and international aerospace organizations. The program structures and methods of conjunction assessment and evasion maneuver systems of aerospace organizations are different. But as developed by inherent principles of the conjunction assessment and evasion maneuver, they share consistent basic frameworks and general procedures.

- (1) Conjunction assessment and collision avoidance maneuver systems of major space agencies.

A. The United States

The conjunction assessment for the U.S. space shuttle in an early stage was adopted with the Box method [18]. To get rid of errors caused by using the Box method, the National Aeronautics and Space Administration, (NASA) started to apply P_c based method in the early 1990s. The conjunction assessments for manned spacecrafts and unmanned spacecrafts of NASA were performed and managed by its subordinate units—Johnson Space Center (JSC) and Goddard Space Flight Center (GSFC).

With regard to manned spacecrafts, NASA has been trying to conduct research on protecting manned spacecraft from collision of traceable space debris since the accident of the space shuttle Challenger in the late 1980s. Thus, JSC was assigned to provide manned spacecraft with conjunction assessment service by using the

Box method and P_c method. With common efforts contributed by NASA JSC and United States Strategic Command (USSTRATCOM), probability methods were studied and relevant software was developed. Moreover, USSTRATCOM provided the data of space station and risk objects while JSC contributed the orbital information of space shuttle. USSTRATCOM was also engaged in debris cataloging and performing procedure to determine the close approaching from space shuttle to space object. In terms of approaching event detection, USSTRATCOM would send the state vector and covariance information to JSC who could calculate and assess the collision probability. For space stations and space shuttles, the red limit and yellow limit of probability for collision avoidance maneuver stipulated by the NASA is 10^{-4} and 10^{-5} , respectively. When the collision probability exceeds the red limit, collision avoidance maneuvers must be conducted. If the probability exceeds the yellow limit, collision avoidance maneuver is still required on the premise of causing no jeopardy to the task object [18].

With regard to unmanned spacecraft, GSFC established a collision risk assessment system of three satellites (Terra, Aqua and Aura) of the Earth Observing System (EOS) for earth science mission in 2004 along with implementation of conventional conjunction assessment procedure [19]. The flight dynamics analysis division of GSFC, in cooperation with the first space control center of USSTRATCOM, provided EOS with the conjunction assessment service. The space control center was for forecasting approaches between EOS satellites and space objects, and the collision risk assessment team of NASA GSFC was in charge of conducting collision risk assessment and developing risk mitigation measures with EOS program. In October, 2005 and June, 2008, EOS's satellites Terra and Aura performed two collision avoidance maneuvers successfully in respective which evaded the debris from risk objects Scout G-1 and OPS-4682 [18], respectively.

In July, 2006, the Joint Space Operations Center (JSpOC) subordinated to USSTRATCOM started to take charge of the Space Object Conjunction Assessment by integrating all Space Surveillance systems in the world, cataloging about 1000 normal-working spacecrafts maintained at high-precision space object orbits and performing forecasting of approaching events. The orbital data of major space objects not only could be retrieved from catalog database, but also could be obtained from the orbital information provided by the spacecraft owner. The number of days which can be forecasted for low earth orbit (LEO) satellites was 5 days and that for geostationary earth orbit (GEO) satellites is 7 days. For GEO and middle earth orbit (MEO) satellites, the reporting rule was when the approaching distance is shortened to be less than 5 km, and the reporting rule for LEO satellites was when the approaching distance is shortened to be less than 1 km.

Conjunction Summary Message (CSM) is shared between JSpOC and satellite owner or manipulator. CSM will be sent from JSpOC to the satellite owner or manipulator via email 72 h prior to TCA to notify the basic information about approaching event, including main object and risk object, TCA, orbital error and components. The satellite owner or manipulator is able to estimate the collision probability by reference to the CSM and develop collision avoidance maneuver

tactics. GSFC's collision assessment and risk analysis team, like many other organizations in the world, makes quantitative analysis by using the approaching event forecasting information given by JSpOC and develops collision avoidance maneuvering measures in cooperation with task administrator, flight dynamics team, and flight operation team.

Due to the aggravation of threats from orbital debris, in August, 2007, NASA requested all LEO and GEO maneuverable spacecrafts to have approaching event assessments on a regular basis to prevent spacecraft collisions with other in-orbit space objects. In April 2009, the applicability of this conventional conjunction assessment analysis was expanded to non-maneuverable spacecrafts and non-working spacecrafts, and the applicability of automatic approaching analytical procedure was expanded to 79 spacecrafts in different orbital regions. In 2009, 8 spacecraft collision avoidance maneuvers were performed by NASA, twice of which were to evade unbroken objects (one was a spacecraft and the other was a rocket body), the rest of collision avoidance maneuvers were performed to evade space debris.

Kelso and Alfano of Center for Space Standards and Innovation (CSSI) established SOCRATES (Satellite Orbital Conjunction Reports Assessing Threatening Encounters in Space) system to provide approaching analysis service and collision risk analysis service [20]. From May 2004, based on the approaching analysis toolbox in STK software, everyday SOCRATES performs approaching analysis on all valid payloads and all in-orbit objects for a following week by taking advantage of the unclassified TLE data and SGP4/SDP4 models and releases the analytical results and used data at CelesTmk website to which users can have free access. Now, the data like approaching time, approaching distance, maximum collision probability, and 10 approaching events with minimum approach distances and maximum collision probabilities for the coming week in respective can be obtained from the website.

B. European Space Agency

The European Space Operations Center (ESOC) of European Space Agency (ESA) has been engaged in observation of the approaching events [21] between its earth remote sensor satellites (i.e., ERS-1, ERS-2, and Envisat) and USSTRATCOM catalog objects since mid-1990s. ESA makes continuous efforts in mathematical method studies and program development and has created the Collision Risk Assessment Tool (CRASS) [22]. This early warning program is able to automatically initiate TLE download, extract object information from database, conduct 7-day approaching event forecast, and report warning & notification information.

At present, ESOC Space Debris Office offers collision risk assessment service to its satellites, including the ERS-2, Envisat and Ciyosat-2 [23]. The main software employed include CRASS and Orbit Determination with Improved Normal equations (ODIN). With regard to ESA's spacecrafts, the acceptable collision probability level is 10^{-4} or an approaching distance less than 300 m. At the end of 2004,

probability thresholds of ERS-2 and Envisat were adjusted to 2×10^{-4} . Satellite Envisat performed collision avoidance maneuver twice successfully in November 2007 and January 2008 to get away from the risk objects of Russian satellites Cosmos-1486 and Cosmos-1624, respectively. In the first quarter of 2010, ESA activated 4 collision avoidance maneuvers, among which three were to evade debris of Iridium-33. In September and October, 2010, ERS-2 and Cryosat-2 made two collision avoidance maneuvers, respectively.

C. Centre National d'Etudes Spatiales

The collision between Satellite Cerise of the Centre National d'Etudes Spatiales (CNES) and debris from upper stage of the rocket Ariane-1 H-10 happened on July 24th, 1997 is proved to be the first accidental collision between two space catalogued objects. Since then, the CNES started to monitor its satellites. Presently, 18 satellites (among which 17 are at LEO and the other is at GEO) are subject to the control of CNES. To avoid accidental collisions, the CNES Orbit Computing Center has taken the responsibility to control the collision risks by using a conjunction assessment program. All data are collected from the data bases of TLE and the GRAVES space object monitoring system in France as well as actual orbit measurements [24]. In the second half of 2009, the CNES forecasted 44 potential dangerous events and took collision avoidance maneuvers for three times. In the first three months of 2010, the CNES forecasted 14 potential dangerous events and took collision avoidance maneuvers for five times.

D. German Space Operations Center

German Space Operations Center (GSOC) started to establish practical conjunction assessment and mitigation system in 2008. The first version of the conjunction event detection software was put into operation in January 2009 and put into practical use in November of that year. Assessment analysis is conducted twice a day to detect the distance between operating satellite and all objects in the TLE catalogue database [25]. Since the establishment of the system, the first collision avoidance maneuver was taken on the TerraSAR-X earth observation satellite of GSOC in November 2009 to avoid the debris from the Russian satellite Cosmos-2251 which was generated in the collision between American satellite Ir-33 and Russian satellite Cosmos-2251.

E. Japan Aerospace Exploration Agency

Japan Aerospace Exploration Agency (JAXA) started to make conjunction assessment on the Advanced Land Observing Satellite (ALOS) in 2008 and had established the framework of collision avoidance maneuvers. Based on the orbit information of operating satellites and space objects catalogued by the SSN, the approaching analysis and collision probability computation can be conducted. The data shall include the data in TLE Catalogue of SSN, ephemeris data of ALOS, and radar-measured data [26]. According to a prediction made in August, 2009, the Rubin4/SL-8 would start to approach ALOS Satellite at a distance of 0.114 km.

The JAXA confirmed the orbit of Rubin4/SL-8 in accordance with the radar-measured data and made an analysis. According to the analysis result, the closest distance between them was 1.596 km. So the analysis was stopped.

(2) General procedures for conjunction assessments and collision avoidance maneuvers

An introduction about the histories, compositions, and functions of the conjunction assessment and collision avoidance maneuver systems of NASA, JSpOC, CSSI, ESA, CNES, GSOC, and JAXA has been made in previous paragraphs. There may be some nuances, but the basic framework and general procedures for conjunction assessment and collision avoidance maneuver are consistent due to the intrinsic rules. Through summaries and analyses, the conjunction assessment and collision avoidance maneuver may involve the following steps:

A. Object filtering and conjunction event prediction. First of all, we shall rule out all objects which are impossible to approach the main object according to the orbit data of the spacecraft (main object) and data in space object catalogue database (generally TLE database of NORAD). Then we shall make a several days (normally 7 days) approaching prediction, determine the approaching time, approaching distance and its components, rendezvous geometrical relationship, relative velocity and other parameters and find out the maximum collision probability. If the approaching distance and the maximum collision probability exceed the pre-set threshold values ($2 \times 25 \times 25$ km for JSpOC [18], 10 km and 10^{-4} for CNES [24], 10^{-4} for GSOC [25], 5 km and 10^{-3} for JAXA [26]), relevant information about the approaching shall be automatically reported once every day.

B. Preliminary assessment and analysis of collision risks. Once a conjunction event is predicted, the collision risk analysis team shall conduct quantitative analysis of the collision risk based on a collision risk assessment system. The quantitative analysis shall involve a calculation of collision probability, an analysis of collision probability sensitivity as well as an evolutionary trend analysis of collision probability. With the latest data and rules including collision probability, orbit error, approaching distance and its components, variation trend, rendezvous geometrical relationship, approaching time, object size, etc., the analysis team shall analyze the dangerous conjunction events by effective methods. Collision risk assessment system is generally composed of modules like an approaching visual module, a 2D collision probability tool module, a 3D nonlinear collision probability tool module, a Monte Carlo collision probability tool module, an evolution and tendency tool module, and other sub-modules.

C. Detailed risk assessments based on observation data. We shall make a detailed assessment of collision risks to find out the dangerous conjunction events which need collision avoidance maneuvers. We shall update all data and determine the orbits of dangerous objects based on radar-measured data to get more accurate analyses. By observing the module or RCS, we will determine the size of dangerous object. All historical TLE data of dangerous objects shall be updated to reduce the

analysis error. Make risk assessment again according to the renewed orbit data to get the latest tendency and statistics analysis results. If the collision probability is still larger than the threshold value, we shall do further analyses.

D. Planning and decisions on collision avoidance maneuver. The collision risk analysis team shall keep doing the risk analysis and work out the most feasible collision avoidance maneuver with the maneuver team. Satellite tasks, platform and operation restrictions as well as the safety of orbit after maneuvers shall be taken into account. Final decision is made based on the evolution of actual collision probability, tendency within the collision scope, task restrictions of the spacecraft, and the competence of performing the collision avoidance maneuver. After deciding to perform the maneuver, we shall also make a decision about whether it is necessary to do the retrograde maneuver. In addition, we shall develop a collision avoidance maneuver strategy, which includes a collision avoidance maneuver scheme and a retrograde maneuver scheme. Firstly, the thrust force can vary within a fixed duration. Secondly, the position of the spacecraft can change along the orbit with a constant velocity increment. The prediction about the orbit after maneuver shall be offered to the collision risk analysis team who can make collision analysis according to the post-maneuver ephemeris to ensure the safety of orbit. Within at least 12 h before the TCA, members concerned shall be called together to make the final decision upon the most feasible collision avoidance maneuver scheme according to the detailed collision risk assessment results.

E. Post-maneuver analysis and assessment. After collision avoidance maneuver, we shall assess the effects of the maneuver and analyze the post-maneuver orbit to ensure the orbit is free from additional risks.

1.3.2 Situational Analysis

Situational analysis is preformed to understand and display complicate system environment. The development tendency in the future can be predicted through the situational analysis.

Situational analysis based on space catalogue data is realized by modeling and analyzing the catalogue data to understand all space situations. Presently, space event detection is one of the hottest topics in the discussion. Along with the advancement in catalogue data research, situational analysis can acquire more information of space objects, which can accordingly deepen the understanding of mankind to space situation.

Space events include space weather events, space object orbit abnormal events, etc. Perception about different space events has become more and more important.

Abnormal space weather may exert serious impact on the satellites on orbit. Solar activities, geomagnetic activities, and other space weather factors have relatively great impacts on the orbit of a space object, and these factors may lead to a sudden change in the orbit energy attenuation rate by changing the upper

atmospheric density. Therefore, the space weather events can be detected according to the abnormal changes of spacecraft orbits if we have a thorough understanding about the impacts of solar activities and geomagnetic activities on the orbit of spacecraft.

Space orbit abnormal events can be used to understand the working conditions of spacecrafts and make corresponding abnormality assessments. Normal spacecraft is able to control its own orbit. To fulfill the missions, the spacecraft shall make a series of actions for the orbital transfer control, orbital maintenance, rendezvous and docking, etc. Therefore, sharp and sudden changes in the orbital elements may occur. A failed spacecraft is uncontrollable and it orbits freely under the influence of gravity and other disturbing forces. The orbital elements change in a stable and continuous way if the spacecraft is intact without any collision, explosion, or disassembly. The working conditions of the spacecraft can be estimated according to the differences of orbit characteristics between normal and abnormal spacecrafts.

References

1. Liou JC (2014) Satellite box score. *Orbital Debris Q News* 18(1):9
2. Hash Y, Bean NP, Steyn WH et al (1998) Attitude control recovery of the CERISE microsatellite following an in-orbit collision. *Adv Astronaut Sci* 98:655–663
3. Johnson N (1996) First natural collision of cataloged earth satellites. *Orbital Debris Q News* 1 (2):1–2
4. Liou JC (2005) Accidental collisions of cataloged satellites identified. *Orbital Debris Q News* 9(2):1
5. Liou JC (2009) Satellite collision leaves significant debris clouds. *Orbital Debris Q News* 13 (2):1
6. Joint chiefs of staff (2013) Space operations. Joint Publication 3–14. May 29th 2013
7. Liu ZG (2000) Design of low orbit space target surveillance net. *J Spacecraft TT&C Technol* 19(4):9–17
8. Qiao K, Wang ZL, Cong MY (2006) Comparison and Analysis on space-based and ground-based surveillance systems of space objects. *Opt Tech* 32(5):744–749
9. Burnham WF, Morton FE, Sridharan R et al (2000) Mission planning for space-based surveillance with the space-based visible sensor. *J Guid Control Dyn* 23(1):165–169
10. Liu L (2000) Orbit theory of spacecraft. National Defense Industry Press, Beijing
11. Hoots FR, Roehrich RL (1980) Space track report no 3: models for propagation of NORAD element sets. Aerospace Defense Command, Peterson, pp 1–79
12. Hoots FR, Schumacher PW, Glover RA (2004) History of analytical orbit modeling in the U. S. space surveillance system. *J Guidance Control Dyn* 27(2):174–185
13. Vallado DA, Crawford P, Hujsak R (2006) Revisiting spacetrack no 3. In: AIAA/AAS astrodynamics specialist conference and exhibit, AIAA 2006-6753, Keystone, Colorado
14. Li JS (2003) Orbit determination of spacecrafts. National Defense Industry Press, Beijing
15. Vallado DA (2005) An analysis of state vector propagation using differing flight dynamics programs, AAS 05-199. In: space flight mechanics conference AAS/AIAA, Copper Mountain, Colorado
16. Fonte DJ (1996) Comparison of orbit propagation in the research and development Goddard trajectory determination system. *Adv Astronaut Sci* 90(2):1–949
17. Wu LD (2011) Orbits and detections of man-made satellites and space debris. Science and Technology Press of China, Beijing

18. Gavin RT (2010) NASA' s orbital debris conjunction assessment and collision avoidance strategy, JSC-CN-19799. NASA Johnson Space Center Flight Dynamics Division, Houston
19. Newman L K, Duncan M (2006) Establishment and implementation of a close approach evaluation and avoidance process for Earth observing system missions. In: astrodynamics specialist conference and exhibit AIAA 2006-6291, AIAA/AAS Aug 21-24 2006, Keystone, Colorado
20. Kelso T S, Alfano S (2005) Satellite orbital conjunction reports assessing threatening encounters in space (SOC-RATES). AAS 05-124. In: 15th AAS/AIAA space flight mechanics conference, Copper Mountain, Colorado
21. Klinkrad H, Alarcon J R, Sanchez N (2005) Collision avoidance for operational ESA satellite. In: Proceedings of the fourth european conference on space debris Darmstadt, Germany
22. Alarcon – Rodriguez JR, Martinez – Fadrique FM, Klinkrad H (2004) Development of a collision risk assessment tool. Adv Space Res 34:1120-1124
23. Flohrer T, Krag H, Klinkrad H (2009) ESA's process for the identification and assessment of high-risk con-junction events. Adv Space Res 44:355-363
24. Laporte F, Sasot E (2008) Operational management of collision risks for LEO satellites at CNES. In: AIAA 2008-3409 SpaceOps 2008 conference, Heidelberg, Germany
25. Aida S, Kirschner M, Wermuth M, et al (2010) Collision avoidance operations for LEO satellites controlled by GSOC. In: AIAA 2010-2298. SpaceOps 2010 conference, Huntsville, Alabama
26. Matsuda I, Hirose C, Kudo N (2010) The JAXA conjunction assessment process. In: AIAA 2010-2039. SpaceOps 2010 conference, Huntsville, Alabama

Chapter 2

Orbital Prediction Error Propagation of Space Objects

2.1 Definition of Coordinate System

In this section, we made a definition and description about all major coordinate systems mentioned in this book.

1. ECI coordinate system

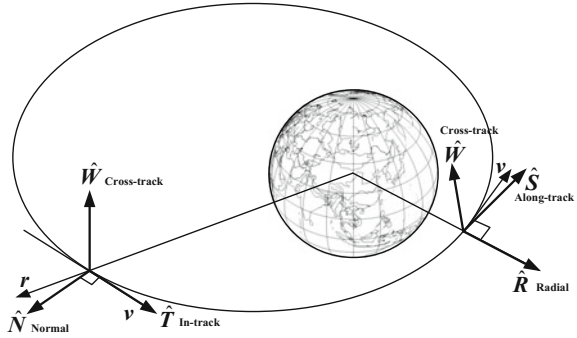
Earth centered inertial (ECI) coordinate system is a common geocentric coordinate system [1, 2]. Its origin is at the center of the earth O_E , and the fundamental plane is the Earth's equator. The X axis is in the equatorial plane, pointing to the epoch mean vernal equinox; the Y axis is in the equatorial plane, at an angle of 90° to the east; the Z axis is vertical to the equatorial plane, coinciding with the earth's spin axis and pointing to the north pole. In view of the time-dependent nature of equator and ecliptic direction under the influence of precession and nutation, the International Astronomical Union (IAU) decided in 1976 to start applying the new epoch standard since 1984. The mean equinox at 12:00:00.00 on January 1st, 2000 has been taken as the benchmark. The coordinate system is also called J2000 ECI coordinate system.

2. Satellite-based orbit coordinate system

Satellite-based orbit coordinate system is the common system used to describe the relative motion of spacecraft, orbit prediction error, and orbital transfer. RSW and NTW coordinate systems [2] are the two common types of the satellite-based orbit coordinate system. Pay attention that there are different ways to define the orbit coordinate system, including the local vertical local horizontal (LVLH) coordinate system, the vehicle velocity local horizontal (VVLH) coordinate system, etc. They are the most common coordinate systems during space rendezvous and docking. In this book, we define the coordinate through Vallado's method [2].

The origin of RSW coordinate system is located at the centroid of the space object. The R axis (radial) always points from the Earth's center along the radius

Fig. 2.1 Satellite-based RSW and NTW coordinate systems



vector toward the satellite as it moves through the orbit; the S axis (along-track or transverse) points in the direction of (but not necessarily parallel to) the velocity vector and is perpendicular to the radius vector; the W axis (cross-track) is normal to the orbital plane. Refer to Fig. 2.1 for details.

Sometimes, the coordinate system is also marked as RTN (radial, transverse, and normal) coordinate system or UVW coordinate system [3]. RSW coordinate system is decided by its position and velocity vectors, which vary with time. Its Axis S does not always coincide with the velocity direction. This kind of coordinate system is mainly used to describe orbital errors, relative position, and satellite orbital transfer of circular orbit.

In an ECI coordinate system, the unit vectors of the three axes of RSW coordinate system are \hat{R} , \hat{S} , and \hat{W} respectively.

$$\hat{R} = \frac{\mathbf{r}}{|\mathbf{r}|}, \quad \hat{W} = \frac{\mathbf{r} \times \mathbf{v}}{|\mathbf{r} \times \mathbf{v}|}, \quad \hat{S} = \hat{W} \times \hat{R} \tag{2.1}$$

The transfer matrix between RSW coordinate system and ECI coordinate system is:

$$\mathbf{M}_{\text{RSW} \rightarrow \text{ECI}} = (\hat{R} \ \hat{S} \ \hat{W}), \quad \mathbf{M}_{\text{ECI} \rightarrow \text{RSW}} = (\hat{R} \ \hat{S} \ \hat{W})^T \tag{2.2}$$

The origin of NTW coordinate system is in the centroid of the space orbit, Axis T (tangential or in-track) is tangent to the orbit and points to the velocity direction, Axis N (normal) is located in the orbital plane, vertical to the velocity direction, positive upward and its Axis W (cross-track) is vertical to the orbital place. Refer to Fig. 2.1 for details. NTW coordinate system also varies with its position and velocity. This kind of coordinate system is mainly used for the atmospheric damping impact analysis and error analysis of elliptical orbits.

In the ECI coordinate system, the unit vectors of the three axes of NTW coordinate system are \hat{N} , \hat{T} , and \hat{W} respectively.

$$\hat{\mathbf{T}} = \frac{\mathbf{v}}{|\mathbf{v}|}, \quad \hat{\mathbf{W}} = \frac{\mathbf{r} \times \mathbf{v}}{|\mathbf{r} \times \mathbf{v}|}, \quad \hat{\mathbf{N}} = \hat{\mathbf{T}} \times \hat{\mathbf{W}} \quad (2.3)$$

The transfer matrix between the NTW coordinate system and the ECI coordinate system is:

$$\mathbf{M}_{\text{NTW} \rightarrow \text{ECI}} = (\hat{\mathbf{N}} \ \hat{\mathbf{T}} \ \hat{\mathbf{W}}), \quad \mathbf{M}_{\text{ECI} \rightarrow \text{NTW}} = (\hat{\mathbf{N}} \ \hat{\mathbf{T}} \ \hat{\mathbf{W}})^{\text{T}} \quad (2.4)$$

NTW coordinate system is also presented as UNW coordinate system. Its Axis U is tangent with orbit and points to velocity direction, Axis N is located in the orbital plane, vertical to the velocity direction, positive upward and its Axis W is vertical to the orbital plane. In the ECI coordinate system, the unit vectors of the three axes of UNW coordinate system are $\hat{\mathbf{U}}$, $\hat{\mathbf{N}}$, and $\hat{\mathbf{W}}$ respectively.

$$\hat{\mathbf{U}} = \frac{\mathbf{v}}{|\mathbf{v}|}, \quad \hat{\mathbf{W}} = \frac{\mathbf{v} \times \mathbf{r}}{|\mathbf{v} \times \mathbf{r}|}, \quad \hat{\mathbf{N}} = \hat{\mathbf{W}} \times \hat{\mathbf{U}} \quad (2.5)$$

$$\mathbf{M}_{\text{UNW} \rightarrow \text{ECI}} = (\hat{\mathbf{U}} \ \hat{\mathbf{N}} \ \hat{\mathbf{W}}), \quad \mathbf{M}_{\text{ECI} \rightarrow \text{UNW}} = (\hat{\mathbf{U}} \ \hat{\mathbf{N}} \ \hat{\mathbf{W}})^{\text{T}}. \quad (2.6)$$

2.2 Overview of Orbital Prediction Errors

2.2.1 Category of Orbital Prediction Errors

The orbital problems mainly include three aspects: Orbital observation, orbital determination, and orbital prediction. Due to various inevitable factors, errors including measuring errors, modeling errors, and methodology errors may exist. The true value of orbit objectively exists, but is unknown. Any orbital measuring method may have errors. An error refers to a deviation and noise. We can only acquire the optimally approximate value of the true orbit value by determining and making assessments with the contaminated data.

Through orbital determination, we can get the initial mean value and initial covariance of the orbit at t_0 . The initial mean value of the orbit is not always equal to the true value. We can get the orbital prediction value and covariance prediction value through orbital propagation. Figure 2.2 is a schematic diagram of the error generation process during space object orbital determination and prediction.

Space object orbital prediction errors include initial errors and model errors. The object state vector at initial epoch time can be acquired through orbital determination after data (radar, photology, GPS, etc.) processing. Due to the errors in measured data, the state vector will inevitably have some errors which are also known as initial state errors. During orbital propagation, the initial state errors will diffuse along with the extrapolation of orbital model and its propagation characteristics and tendency will vary according to the types of orbits. Based on the

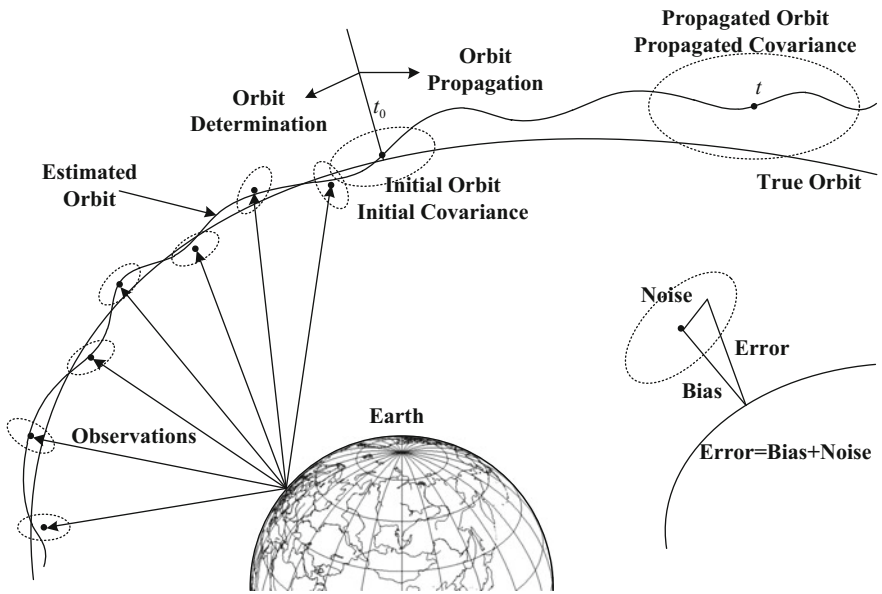


Fig. 2.2 Schematic diagram of errors during space object orbital determination and prediction

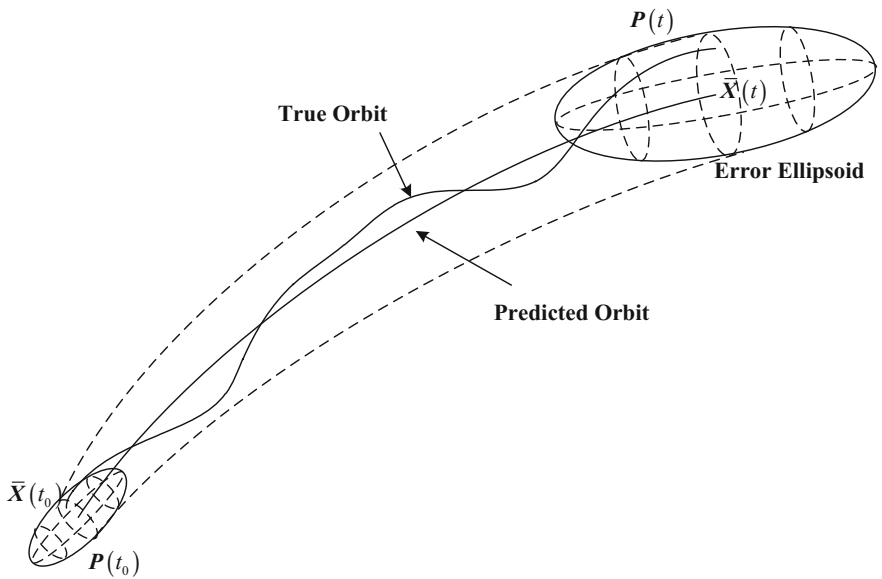


Fig. 2.3 Propagation of initial errors

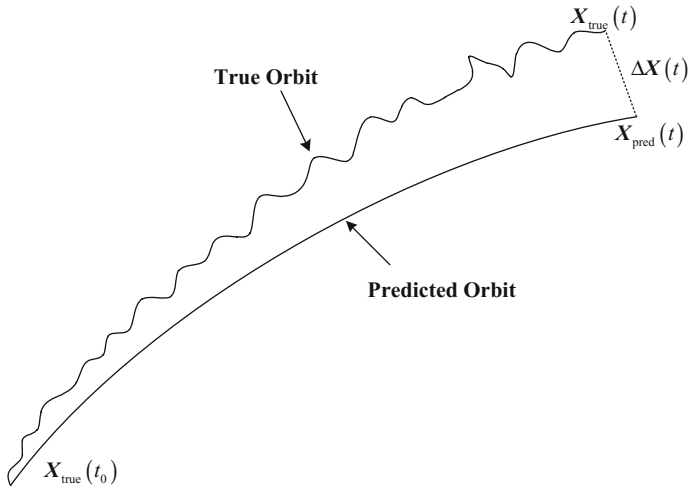


Fig. 2.4 Orbital model errors

propagation of orbital mean value and covariance, we can work out an error pipe in which the orbital true value exists at a high probability. Refer to Fig. 2.3.

The incompleteness of orbital propagation model can cause model errors during orbital propagation. If we can get the orbital true value at the initial time, model errors may bring about orbital propagation errors, refer to Fig. 2.4. The diffusion of initial errors and model errors may be coupled with each other. Orbital prediction error analysis is a complicated problem.

During orbital propagation, the magnitudes of initial errors are different from those of model errors. For the orbital propagation model with a relatively high precision, its initial errors will be much greater than the model errors. To simplify the analysis process, the analysis of initial errors is decoupled with that of the model errors. When analyzing initial errors, the orbital prediction model shall be regarded as correct without any model error and the propagation characteristics of initial errors will be obtained. When analyzing model errors, different models with the same initial state shall be taken for the orbital propagation, and the propagation results shall be compared with each other.

One of the major problems for object monitoring and space object orbital data application is acquiring the orbital propagation error covariance information of space object and knowing the precision of that information. In practice, we usually use the inner fitting precision and outer comparing precisions to define the precision of orbital determination [4]. Considering the two assessment methods and the actual engineering conditions, the analytical methods elaborated in the domestic and foreign literature are error extrapolation based on orbital models, comparison with high-precision orbital prediction results, and error analysis based on historical

orbital data in allusion to different objects and sources of their orbit information. In this chapter, we have a brief introduction about the first two methods. And the third method will be discussed in the next chapter.

2.2.2 Model-Based Orbital Error Propagation

For the space cooperative objects like the normal spacecraft owned or controlled by one's own side, the owner or operator of the spacecraft shall take the responsibility for the orbital determination, acquire the measured data with a relatively high precision by various means, obtain the orbital elements and motion state vector based on the orbital determination theories, and get the error information of orbital determination or initial covariance matrix. Initial covariance matrix propagation can be realized with transfer matrix which depends on the orbital model we used.

Due to the nonlinear dynamic orbit model, we cannot get the analytical transfer matrix. In addition to the numerical analysis method, covariance analysis describing function technique (CADET) is a common tool for linearization error analysis. The Analytic Sciences Corporation of U.S. put forward CADET to analyze the statistical properties of linear or nonlinear systems with random errors. This method is featured with the characteristics of time saving, high efficiency, and high reliability [5–8]. The basic concept of CADET is: firstly, we conduct the statistical linearization on the nonlinear system based on the describing function theory; then, work out the mean values of the random state variables and the covariance propagation differential equation of the linearized system based on CADET. Once the two differential equations are worked out, we can determine the statistical properties of system state variables through one-time solution. Therefore, compared with the sampling statistics method (Monte Carlo Method), CADET can save more time. In addition, we can get accurate solutions for linear system by applying CADET. For nonlinear systems, approximate values can also be obtained by means of statistical linearization.

Liang et al. put forward the CADET method for error propagation analysis [9, 10] in allusion to the rendezvous and docking issues. In view of the navigation and control errors, CADET propagation model for relative states and deviation statistical variables has been established based on the C-W equation. In view of the orbital perturbation model, navigation and control errors, the CADET propagation model for relative states and deviation statistical variables has been established by nonlinear perturbation dynamical equation.

We can get the error covariance matrix with a high precision through an orbital error propagation, which requires the covariance at initial time and orbit model with high precision. However, this is difficult to satisfy under normal circumstances. This method takes a long time for calculation and cannot be applied to a large batch of objects, so it is only applicable to some controllable cooperative objects. In general cases, we usually adopt the method for the generation of orbit propagation errors of key spacecrafts in space surveillance missions.

2.2.3 Comparison with High Precision Ephemeris

The orbital prediction precision can be evaluated by comparing the orbital prediction ephemeris with high precision ephemeris. Higher precision orbital data for reference are usually obtained from the owner of spacecraft, orbital determination observation, or data from the high precision orbital predictor. Vallado has done a relatively thorough comparison about the difference of each orbital prediction model [11, 12]. Chan compared the prediction results [13] of TLE prediction orbit and double-station ranging orbital determination system of Intelsat's geostationary orbital communication satellite group, analyzed the variation of errors with time, and gave out the different characteristics of errors over different marine areas. Kelso analyzed the TLE orbital prediction results and the high precision ephemeris of GPS satellites, as well as the consistency of TLE data [14]. Boyce also analyzed the TLE prediction errors of Iridium constellation [15]. Snow compared the precisions of TLE data and the corresponding measured data that are used to generate the TLE data [16]. Muldoon improved the orbital determination and propagation with GPS measured data [17]. Before and after the epoch time of TLE, estimated initial error at epoch time can be obtained by comparing the high precision orbital determination results with the TLE data to get the initial value for error propagation of the orbit.

For most users, this precision assessment method which can only be used for the result comparison and verification is inapplicable, as it is impossible to get the high precision orbital determination results of most objects in space catalogue and most users are incapable to get the measured data or high precision ephemeris. In addition, the orbital prediction error of each object is unique, requiring separate analysis. Few objects with high precision ephemeris cannot represent all objects on the same orbit completely.

2.3 Covariance Analysis Describing Function Technique

2.3.1 Covariance Analysis of Linear System

The differential equation of first order can be applied to the description of random system in a linear continuous time domain:

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{G}(t)\mathbf{w}(t) + \mathbf{D}(t) \quad (2.7)$$

Wherein: $\mathbf{x}(t)$ is the state vector of n-dimensional systems, which is composed of plenty of state variables that can describe the system state; $\mathbf{w}(t)$ is the m-dimensional disturbance and also the stochastic noise vector of the control input of the system; $\mathbf{F}(t)$ and $\mathbf{G}(t)$ are deterministic functional matrixes and $\mathbf{D}(t)$ is deterministic functional vector.

As random state $\mathbf{x}(t)$ is made up of mean value $\bar{\mathbf{x}}(t)$ and random component $\delta\mathbf{x}(t)$, the random system in a linear continuous time domain can be described as follow:

$$\mathbf{x}(t) = \bar{\mathbf{x}}(t) + \delta\mathbf{x}(t) \quad (2.8)$$

$$\bar{\mathbf{x}}(t) = E\{\mathbf{x}(t)\}, \quad \mathbf{P}(t) = E\{\delta\mathbf{x}(t)\delta\mathbf{x}^T(t)\} \quad (2.9)$$

Similarly, random noise vector $\mathbf{w}(t)$ is composed of mean value $\bar{\mathbf{w}}(t)$ and random component $\delta\mathbf{w}(t)$, the random system in a linear continuous time domain can be described as follows:

$$\mathbf{w}(t) = \bar{\mathbf{w}}(t) + \delta\mathbf{w}(t) \quad (2.10)$$

$$\bar{\mathbf{w}}(t) = E\{\mathbf{w}(t)\} \quad (2.11)$$

When the disturbance is white noise, the covariance matrix of process noise can be presented as:

$$E\{\delta\mathbf{w}(\tau)\delta\mathbf{w}^T(t)\} = \mathbf{Q}(t)\delta(t - \tau) \quad (2.12)$$

Wherein, $\delta(t - \tau)$ is the Dirac function, which has the following properties:

$$\delta(t - \tau) = \begin{cases} 0 & (t \neq \tau) \\ \infty & (t = \tau) \end{cases}, \quad \int_{-\infty}^{\infty} \delta(\zeta)d\zeta = 1 \quad (2.13)$$

According to the describing function theory of linear system, the equation set of mean value and covariance propagation is shown as follows:

$$\begin{cases} \dot{\bar{\mathbf{x}}}(t) = \mathbf{F}(t)\bar{\mathbf{x}}(t) + \mathbf{G}(t)\bar{\mathbf{w}}(t) + \mathbf{D}(t) \\ \dot{\mathbf{P}}(t) = \mathbf{F}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}^T(t) + \mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}^T(t) \end{cases} \quad (2.14)$$

Wherein: the diagonal elements of covariance matrix are the mean square errors of the random parts of state variables and the off-diagonal elements describe the correlation among all state variables. The covariance matrix is composed of these two kinds of elements. Based on the mean value and the covariance propagation equation set, we can work out the statistical rules of system state variables. So we can get the state mean and variance at any time by knowing the initial conditions $\bar{\mathbf{x}}(t_0)$ and $\mathbf{P}(t_0)$. Second item of Eq. (2.14) is same as the linear system matrix Ricatti Equation [8].

2.3.2 Covariance Analysis of Nonlinear System

The differential equation for the nonlinear system in a continuous time domain:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}, t) + \mathbf{G}(t)\mathbf{w}(t) + \mathbf{D}(t) \quad (2.15)$$

where $\mathbf{f}(\mathbf{x}, t)$ is a nonlinear vector function of the n -dimensional state variable $\mathbf{x}(t)$. Definitions of other functions are the same as the definitions in linear system.

For CADET method for nonlinear system, first of all, we shall conduct the statistical linearization on the nonlinear system by applying the describing function theory. Then, we need to analyze the statistic performance of the linear system by using the covariance analysis method.

With the purpose of turning the nonlinear function $\mathbf{f}(\mathbf{x}, t)$ into the approximate linear function of state variable $\mathbf{x}(t)$, the linearization includes true linearization and quasi-linearization. Taylor series expansion method is the most common method in true linearization. Statistical linearization belongs to quasi-linearization. According to the probability density function form of state function $\mathbf{x}(t)$, statistical linearization is to use a linear function $\mathbf{f}(\mathbf{x})$ to approximate the nonlinear function within a large variation range of $\mathbf{x}(t)$. The nonlinear function $\mathbf{f}(\mathbf{x})$ is not required to be continuous and differentiable. This method makes the linearization of a large number of nonlinear functions possible, such as the step, saturation control, and numerical switch functions that are frequently used in flight vehicles, which cannot be linearized by using true linearization methods. Statistic linearization, however, has disadvantages. We have to know the form of the probability density function of $\mathbf{x}(t)$. Therefore, the same function with different probability density will have different linearization results.

Statistic linearization steps: for nonlinear vector function $\mathbf{f}(\mathbf{x}, t)$ of random state variable $\mathbf{x}(t)$, work out the quasi-linear equation $\hat{\mathbf{f}} + N_{\delta\mathbf{x}}\delta\mathbf{x}$ to minimize the mean square error $E\{\mathbf{e}^T\mathbf{e}\}$ between this equation and vector function $\mathbf{f}(\mathbf{x}, t)$. Wherein:

$$\mathbf{e} = \mathbf{f} - N_{\bar{\mathbf{x}}}\bar{\mathbf{x}} - N_{\delta\mathbf{x}}\delta\mathbf{x} = \mathbf{f} - \hat{\mathbf{f}} - N_{\delta\mathbf{x}}\delta\mathbf{x} \quad (2.16)$$

Shall meet the:

$$\frac{\partial}{\partial N_{\bar{\mathbf{x}}}} \text{tr}(E\{\mathbf{e}^T\mathbf{e}\}) = \frac{\partial}{\partial N_{\delta\mathbf{x}}} \text{trace}(E\{\mathbf{e}^T\mathbf{e}\}) = 0 \quad (2.17)$$

Assume

$$N_{\bar{\mathbf{x}}}(\bar{\mathbf{x}}, \mathbf{P})\bar{\mathbf{x}} = E\{\mathbf{f}(\mathbf{x}, t)\} \quad (2.18)$$

The describing function can be obtained,

$$\begin{aligned}\hat{\mathbf{f}}(t) &= E\{\mathbf{f}(\mathbf{x}, t)\} \\ \mathbf{N}_{\delta\mathbf{x}}(t) &= E\{\mathbf{f}(\mathbf{x}, t)\delta\mathbf{x}^T(t)\}\mathbf{P}^{-1}(t)\end{aligned}\quad (2.19)$$

where $\mathbf{N}_{\delta\mathbf{x}}$ is known as a dynamic matrix of quasi-linear system; $\hat{\mathbf{f}}$ is the expectant vector which can be formulated as:

$$\hat{\mathbf{f}} = E\{\mathbf{f}(\mathbf{x}, t)\} = \int_{-\infty}^{\infty} \mathbf{f}(\mathbf{x}, t)p(\mathbf{x})d\mathbf{x}\quad (2.20)$$

where $p(\mathbf{x})$ is the probability density function of $\mathbf{x}(t)$.

Statistic linearization theory shall be applied, i.e.,

$$\mathbf{f} \simeq \hat{\mathbf{f}} + \mathbf{N}_{\delta\mathbf{x}}\delta\mathbf{x}\quad (2.21)$$

Substitute the above equation into Eq. (2.15), we can get:

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \hat{\mathbf{f}} + \mathbf{N}_{\delta\mathbf{x}}\delta\mathbf{x} + \mathbf{G}(t)\mathbf{w}(t) + \mathbf{D}(t) \\ &= \mathbf{N}_{\delta\mathbf{x}}\mathbf{x}(t) + \mathbf{G}(t)\mathbf{w}(t) + \mathbf{R}(t)\end{aligned}\quad (2.22)$$

where

$$\mathbf{R}(t) = \hat{\mathbf{f}} - \mathbf{N}_{\delta\mathbf{x}}\bar{\mathbf{x}}(t) + \mathbf{D}(t)\quad (2.23)$$

where both $\mathbf{N}_{\delta\mathbf{x}}$ and $\hat{\mathbf{f}}$ do not contain $\mathbf{x}(t)$, and:

$$\mathbf{x}(t) = \bar{\mathbf{x}}(t) + \delta\mathbf{x}(t)\quad (2.24)$$

$$\bar{\mathbf{x}}(t) = E\{\mathbf{x}(t)\}\quad (2.25)$$

$$\mathbf{P}(t) = E\{\delta\mathbf{x}(t)\delta\mathbf{x}^T(t)\}\quad (2.26)$$

According to covariance analysis theory for linear system, the propagation equations of mean vector $\bar{\mathbf{x}}(t)$ and covariance matrix $\mathbf{P}(t)$ of the random state vector $\mathbf{x}(t)$ under the disturbance of which noise is shown as follows:

$$\begin{cases} \dot{\bar{\mathbf{x}}}(t) = \hat{\mathbf{f}}(t) + \mathbf{G}(t)\bar{\mathbf{w}}(t) + \mathbf{D}(t) \\ \dot{\mathbf{P}}(t) = \mathbf{N}_{\delta\mathbf{x}}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{N}_{\delta\mathbf{x}}^T(t) + \mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}^T(t) \end{cases}\quad (2.27)$$

In general, $\hat{\mathbf{f}}(t)$ and $\mathbf{N}_{\delta\mathbf{x}}$ contain the mean value $\bar{\mathbf{x}}(t)$ and covariance $\mathbf{P}(t)$ of $\mathbf{x}(t)$, the propagation equations above are nonlinear equations.

2.3.3 Gaussian Distribution-Based Describing Function

Most of the common random state vectors are subject to Gaussian distribution, so the state of a space object can also be regarded to be subject to Gaussian distribution. In this section, we firstly obtain the statistic linear describing function of nonlinear system when random variable $\mathbf{x}(t)$ is subject to joint Gaussian distribution. Then, apply the result to the orbital dynamic equation to analyze the propagation situation of the estimated mean and estimated covariance in high-precision orbital prediction.

According to the definition of describing function, in order to work out the describing function, we have to acquire the probability density function of $\mathbf{x}(t)$. In general, the analytic representations are complicate and even impossible to be obtained. If $\mathbf{x}(t)$ is subject to joint Gaussian distribution, the probability density function will be:

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{P}|}} \exp\left(-\frac{1}{2} \delta \mathbf{x}^T \mathbf{P}^{-1} \delta \mathbf{x}\right) \quad (2.28)$$

The describing function calculation can be simplified as follows:

$$\begin{aligned} \hat{\mathbf{f}} &= E\{\mathbf{f}(\mathbf{x}, t)\} = \int_{-\infty}^{\infty} \mathbf{f}(\mathbf{x}, t) p(\mathbf{x}) d\mathbf{x} \\ N_{\delta \mathbf{x}} &= \frac{d\hat{\mathbf{f}}}{d\bar{\mathbf{x}}} \end{aligned} \quad (2.29)$$

namely:

$$\hat{\mathbf{f}}(t) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{P}|}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathbf{f}(\mathbf{x}, t) \exp\left(-\frac{1}{2} \delta \mathbf{x}^T \mathbf{P}^{-1} \delta \mathbf{x}\right) dx_1 \cdots dx_n \quad (2.30)$$

$$\begin{aligned} \frac{d\hat{\mathbf{f}}}{d\bar{\mathbf{x}}} &\simeq \left(\frac{\partial \hat{\mathbf{f}}}{\partial x_1} \quad \cdots \quad \frac{\partial \hat{\mathbf{f}}}{\partial x_n} \right)^T \\ &= \frac{1}{\sqrt{(2\pi)^n |\mathbf{P}|}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathbf{f}(\mathbf{x}, t) \exp\left(-\frac{1}{2} \delta \mathbf{x}^T \mathbf{P}^{-1} \delta \mathbf{x}\right) \delta \mathbf{x}^T dx_1 \cdots dx_n \mathbf{P}^{-1} \\ &= E\{\mathbf{f}(\mathbf{x}, t) \delta \mathbf{x}^T\} \mathbf{P}^{-1} \end{aligned} \quad (2.31)$$

According to the definition of the dynamic matrix describing function of the quasi-linear system, we can learn that:

$$\mathbf{N}_{\delta\mathbf{x}} = \frac{d\hat{\mathbf{f}}}{d\bar{\mathbf{x}}} \quad (2.32)$$

Now, the concrete solution to describing functions is presented. Set the statistic linear describing function of state equation $\mathbf{f} = (f_1, \dots, f_n)^T$ as $\hat{\mathbf{f}} = (\hat{f}_1, \dots, \hat{f}_n)^T$. If the continuous function vector $\mathbf{f}(\mathbf{x})$ of the random variable has no sine/cosine function, the Taylor expansion at mean value $\bar{\mathbf{x}}$ shall be presented as follows: (the case with sine/cosine functions shall be subject to special treatment, refer to literature [6])

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\bar{\mathbf{x}} + \delta\mathbf{x}) = \mathbf{f}(\bar{\mathbf{x}}) + \mathbf{D}_{\delta\mathbf{x}}\mathbf{f} + \frac{1}{2}\mathbf{D}_{\delta\mathbf{x}}^2\mathbf{f} + \mathbf{R}_2(\mathbf{x}) \quad (2.33)$$

where $\mathbf{R}_2(\mathbf{x})$ is the remainder of Taylor Formula in second order; $\mathbf{D}_{\delta\mathbf{x}}\mathbf{f}$ is the total differential of $\mathbf{f}(\cdot)$ disturbed at $\bar{\mathbf{x}}$ caused by disturbance $\delta\mathbf{x}$. The equations in the first order and second order shall be presented respectively as follows:

$$\begin{cases} \mathbf{D}_{\delta\mathbf{x}}\mathbf{f} = \sum_{p=1}^n (x_p - \bar{x}_p) \frac{\partial}{\partial x_p} \mathbf{f}(\mathbf{x}) \Big|_{\mathbf{x}=\bar{\mathbf{x}}} \\ \mathbf{D}_{\delta\mathbf{x}}^2\mathbf{f} = \sum_{p=1}^n \sum_{q=1}^n (x_p - \bar{x}_p)(x_q - \bar{x}_q) \frac{\partial^2}{\partial x_p \partial x_q} \mathbf{f}(\mathbf{x}) \Big|_{\mathbf{x}=\bar{\mathbf{x}}} \end{cases} \quad (2.34)$$

Ignoring $\mathbf{R}_2(\mathbf{x})$ and doing the expectation of Eq. (2.33), the function vector $\mathbf{f}(\mathbf{x})$ can be linearized as follows:

$$E\{\mathbf{f}(\mathbf{x})\} = E\{\mathbf{f}(\bar{\mathbf{x}})\} + E\{\mathbf{D}_{\delta\mathbf{x}}\mathbf{f}\} + \frac{1}{2}E\{\mathbf{D}_{\delta\mathbf{x}}^2\mathbf{f}\} \quad (2.35)$$

The scalar form is

$$\begin{aligned} E\{f(\mathbf{x})\} &= E\{f(\bar{\mathbf{x}})\} + E\left\{ \sum_{p=1}^n (x_p - \bar{x}_p) \frac{\partial}{\partial x_p} f(\mathbf{x}) \Big|_{\mathbf{x}=\bar{\mathbf{x}}} \right\} \\ &+ \frac{1}{2}E\left\{ \sum_{p=1}^n \sum_{q=1}^n (x_p - \bar{x}_p)(x_q - \bar{x}_q) \frac{\partial^2}{\partial x_p \partial x_q} f(\mathbf{x}) \Big|_{\mathbf{x}=\bar{\mathbf{x}}} \right\} \end{aligned} \quad (2.36)$$

It can be noticed that:

$$E\{\mathbf{D}_{\delta\mathbf{x}}\mathbf{f}\} = E\left\{ \sum_{p=1}^n (x_p - \bar{x}_p) \frac{\partial}{\partial x_p} \mathbf{f}(\mathbf{x}) \Big|_{\mathbf{x}=\bar{\mathbf{x}}} \right\} = 0 \quad (2.37)$$

Then,

$$E\{f(\mathbf{x})\} = f(\bar{\mathbf{x}}) + \frac{1}{2}E\left\{\sum_{p=1}^n \sum_{q=1}^n (x_p - \bar{x}_p)(x_q - \bar{x}_q) \frac{\partial^2}{\partial x_p \partial x_q} f(\mathbf{x}) \Big|_{\mathbf{x}=\bar{\mathbf{x}}}\right\} \quad (2.38)$$

According to the statistic theory, the statistic linear describing function of $f(\mathbf{x})$ is:

$$\hat{f}(\mathbf{x}) = E\{f(\mathbf{x})\} = f(\bar{\mathbf{x}}) + \frac{1}{2}E\left\{\sum_{p=1}^n \sum_{q=1}^n (x_p - \bar{x}_p)(x_q - \bar{x}_q) \frac{\partial^2}{\partial x_p \partial x_q} f(\mathbf{x}) \Big|_{\mathbf{x}=\bar{\mathbf{x}}}\right\} \quad (2.39)$$

Where the i th factor of $\hat{f}(\mathbf{x})$ shall be:

$$\hat{f}_i(\mathbf{x}) = f_i(\bar{\mathbf{x}}) + \frac{1}{2}E\left\{\sum_{p=1}^n \sum_{q=1}^n (x_p - \bar{x}_p)(x_q - \bar{x}_q) \frac{\partial^2}{\partial x_p \partial x_q} f_i(\mathbf{x}) \Big|_{\mathbf{x}=\bar{\mathbf{x}}}\right\} \quad (2.40)$$

Covariance matrix of vector \mathbf{x} about the mean value $\bar{\mathbf{x}}$ shall be:

$$\begin{aligned} \mathbf{P} &= E\{\delta\mathbf{x}\delta\mathbf{x}^T\} = E\{(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^T\} \\ &= \begin{pmatrix} E\{(x_1 - \bar{x}_1)^2\} & \cdots & E\{(x_1 - \bar{x}_1)(x_n - \bar{x}_n)\} \\ \vdots & & \vdots \\ E\{(x_n - \bar{x}_n)(x_1 - \bar{x}_1)\} & \cdots & E\{(x_n - \bar{x}_n)^2\} \end{pmatrix} \triangleq \begin{pmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & & \vdots \\ P_{n1} & \cdots & P_{nn} \end{pmatrix} \end{aligned} \quad (2.41)$$

Then, Eqs. (2.39) and (2.40) can be presented as follows:

$$\hat{f}(\mathbf{x}) = f(\bar{\mathbf{x}}) + \frac{1}{2} \sum_{p=1}^n \sum_{q=1}^n P_{pq} \frac{\partial^2}{\partial x_p \partial x_q} f(\mathbf{x}) \Big|_{\mathbf{x}=\bar{\mathbf{x}}} \quad (2.42)$$

$$\hat{f}_i(\mathbf{x}) = f_i(\bar{\mathbf{x}}) + \frac{1}{2} \sum_{p=1}^n \sum_{q=1}^n P_{pq} \frac{\partial^2}{\partial x_p \partial x_q} f_i(\mathbf{x}) \Big|_{\mathbf{x}=\bar{\mathbf{x}}} \quad (2.43)$$

At this moment,

$$\mathbf{N}_{\delta\mathbf{x}} = \frac{d\hat{f}}{d\bar{\mathbf{x}}} = \frac{df(\bar{\mathbf{x}})}{d\bar{\mathbf{x}}} + \frac{d}{d\bar{\mathbf{x}}} \left(\frac{1}{2} \sum_{p=1}^n \sum_{q=1}^n P_{pq} \frac{\partial^2}{\partial x_p \partial x_q} f_i(\mathbf{x}) \Big|_{\mathbf{x}=\bar{\mathbf{x}}} \right) \quad (2.44)$$

Comparing with linear system, Eq. (2.44) derived from the quasi-linear system has a second-order corrective term. Therefore, the description on covariance propagation is much more accurate than the Ricatti Formula for nonlinear system.

If the signal and information systems concerned are nonlinear systems, and the optimal estimation is not limited to be a linear one, or the noise is in non-Gaussian distribution, it will be relatively difficult to present the covariance matrix. The martingale representation of nonlinear system error covariance matrix is presented in literature [18]. When doing covariance analysis, high-order statistics approaching method shall be adopted.

Providing that the initial state error is Gaussian white noise, a comparison had been made upon the covariance matrixes which were worked out through CADET and Monte Carlos Method. The comparison in literature [19] presents the same result which has a relatively higher calculating precision.

In addition, literature [20] provides a reliable research on the non-Gaussian error propagation of orbit and substitutes the Gaussian Confidence Ellipsoid with the maximum boundary calibrated error range of undetermined forecasting cloud.

2.3.4 CADET's Application on Orbit

Space object dynamic differential equation can be presented as:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t) + \mathbf{w}(t), \quad \mathbf{w}(t) \sim N(0, \mathbf{Q}(t)) \quad (2.45)$$

Providing that the system noise $\mathbf{w}(t)$ is white noise with an expectation value of 0, the spectrum density matrix is $\mathbf{Q}(t)$. And the orbital state estimated mean value and covariance at initial time shall be:

$$\mathbf{x}(0) \sim N(\bar{\mathbf{x}}_0, \mathbf{P}_0) \quad (2.46)$$

According to the covariance describing function method, orbital propagation is to integrate the following propagation equations of estimated mean value and covariance:

$$\dot{\hat{\mathbf{x}}}(t) = \hat{\mathbf{f}}(\bar{\mathbf{x}}(t), t) + \bar{\mathbf{w}}(t) \quad (2.47)$$

$$\hat{\mathbf{f}}(\bar{\mathbf{x}}(t), t) = \mathbf{f}(\bar{\mathbf{x}}(t), t) + \frac{1}{2} \sum_{p=1}^n \sum_{q=1}^n P_{pq} \frac{\partial^2}{\partial x_p \partial x_q} \mathbf{f}(\mathbf{x}(t), t) \Big|_{\mathbf{x}=\bar{\mathbf{x}}} \quad (2.48)$$

$$\dot{\mathbf{P}}(t) = \mathbf{F}(\bar{\mathbf{x}}(t), t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}^T(\bar{\mathbf{x}}(t), t) + \mathbf{Q}(t) \quad (2.49)$$

$$\mathbf{F}(\bar{\mathbf{x}}(t), t) = \frac{\partial \hat{\mathbf{f}}(\bar{\mathbf{x}}(t), t)}{\partial \bar{\mathbf{x}}(t)} \quad (2.50)$$

Providing that the mean value of state vector \mathbf{x} is $\bar{\mathbf{x}} = (\bar{\mathbf{r}}^T, \bar{\mathbf{v}}^T)^T$, the differential equation of mean value shall be $\dot{\bar{\mathbf{x}}} = (\bar{\mathbf{v}}^T, \bar{\mathbf{a}}^T)^T$. If the geocentric distance is 7000 km, the second-order partial derivative of the acceleration of gravity with respect to the radius vector is with the magnitude of $d^2(\mu/r^2)/dr^2 \approx 10^{-12}$.

If the position variance is with a magnitude of 10^6 m^2 , the magnitude of the second-order term of Taylor expansion is about 10^{-6} , which is equal to the magnitude of the non-spherical perturbation term J_4 . In general calculation, only the J_2 impact of the compression of the Earth is considered in the second-order partial derivative. If the orbital altitude is relatively low, the second-order partial derivative of atmospheric damping shall also be taken into account.

Integral of mean value and covariance shall be conducted in ECI coordinate system and error analysis shall be conducted in RSW coordinate system as follows:

$$\mathbf{P}_{\text{RSW}} = \mathbf{J}\mathbf{P}_{\text{ECI}}\mathbf{J}^T \quad (2.51)$$

$$\mathbf{J} = \begin{pmatrix} \mathbf{M}_{\text{RSW} \rightarrow \text{ECI}}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\text{RSW} \rightarrow \text{ECI}}^T \end{pmatrix} = \begin{pmatrix} (\hat{\mathbf{R}} \ \hat{\mathbf{S}} \ \hat{\mathbf{W}})^T & \mathbf{0} \\ \mathbf{0} & (\hat{\mathbf{R}} \ \hat{\mathbf{S}} \ \hat{\mathbf{W}})^T \end{pmatrix} \quad (2.52)$$

According to the space object dynamics equation mentioned in last section, we can analysis the initial mean value and error propagation situation in this section. Ignoring the impact of noise and considering the impact of major second-order partial derivatives, at this moment the mean value differential equation will be the same as Eq. (2.47) and the covariance propagation equation is:

$$\dot{\mathbf{P}}(t) = \mathbf{F}(\bar{\mathbf{x}}(t), t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}^T(\bar{\mathbf{x}}(t), t) \quad (2.53)$$

The initial condition is

$$\mathbf{x}(0) \sim N(\bar{\mathbf{x}}_0, \mathbf{P}_0) \quad (2.54)$$

Select four typical space objects including the Sun Synchronous Orbit (SSO), the MEO, the Highly Eccentric Orbit (HEO), and the GEO. Refer to Table 2.1 for the types of objects and orbits. Refer to Table 2.2 for the orbital elements of objects (epoch times are UTC 00:00:00, August 1st 2008).

Table 2.1 Types of objects and orbits

S/N	Type of orbit	Type of object	Object name	Object code
①	SSO	Earth Resources Satellite	ERS-2	23560
②	MEO	GPS Satellite	GPSBlockHa	25030
③	HEO	Molnyia Satellite	Molnyia3-35	20052
④	GEO	Galaxy Communication Satellite	Galaxy-11	26038

Table 2.2 Orbital elements of objects

S/N	a/km	e	$i/(\text{°})$	$\Omega/(\text{°})$	$\omega/(\text{°})$	$M/(\text{°})$
①	7151.800091	0.0000917	98.5690	76.5485	99.0473	261.0823
②	26558.682691	0.010179	56.277	62.334	163.339	273.257
③	25663.740026	0.740361	63.463	359.792	270.021	40.168
④	42166.577215	0.000036	0.068	109.622	217.823	41.434

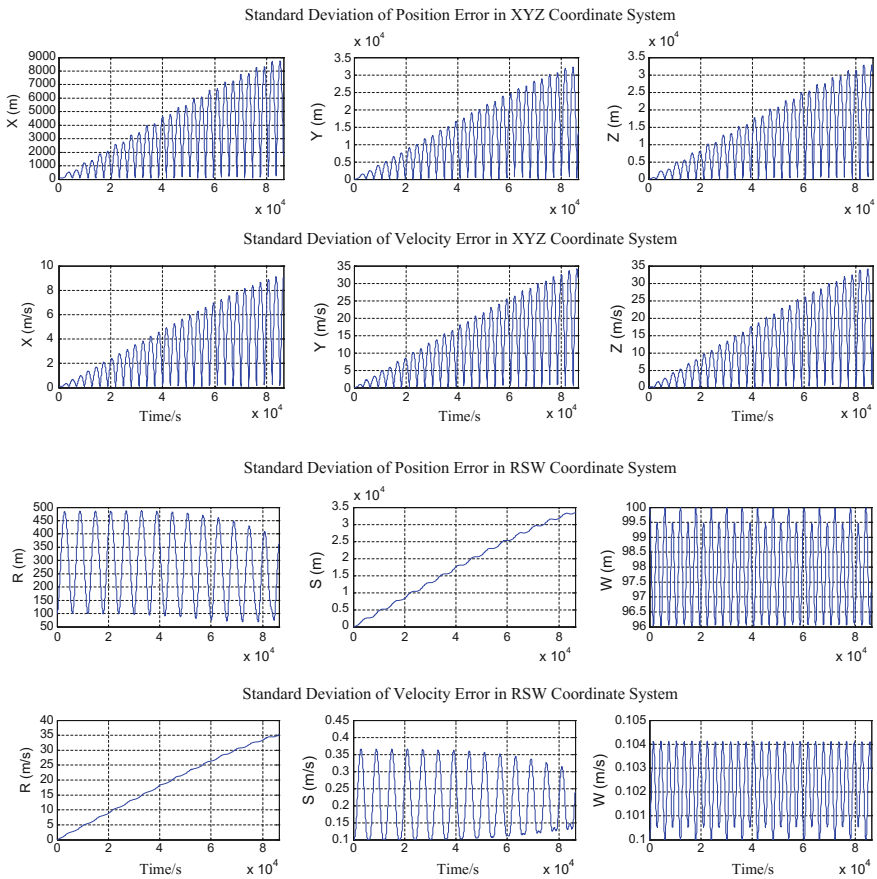


Fig. 2.5 Error propagation curve of object ①

Ignoring the model noise and providing that the standard deviations of the position error in each axis direction at initial state is 100 m and that of the velocity error is 0.1 m, the error divergence situations in XYZ and RSW coordination systems in the next day are shown as Figs. 2.5 2.6, 2.7, 2.8 and 2.9.

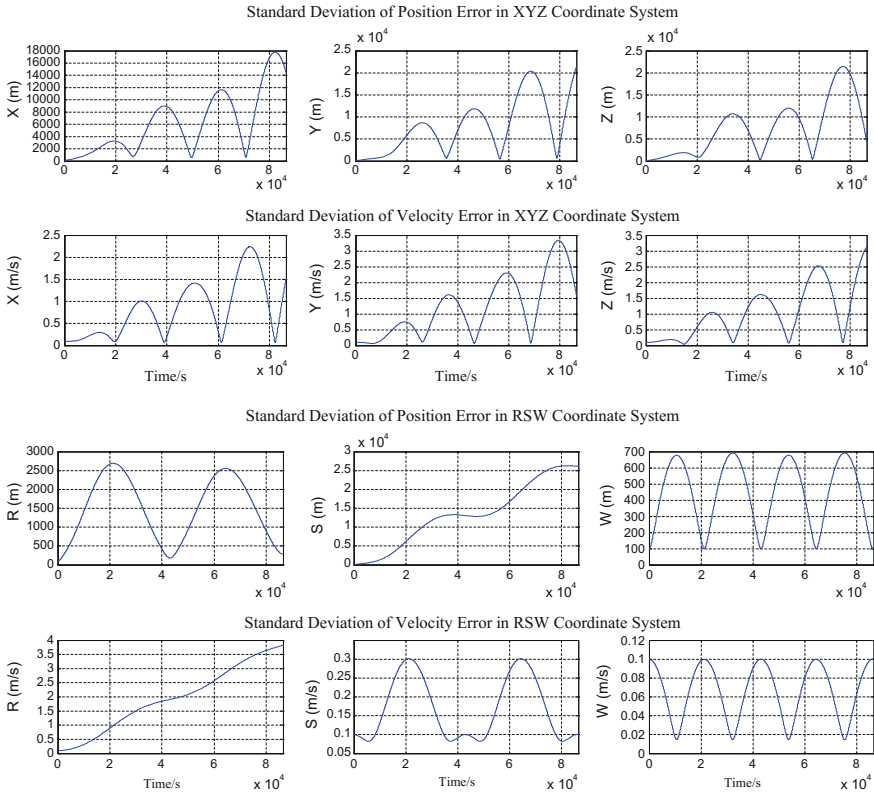


Fig. 2.6 Error propagation curve of object ②

As shown in Figs. 2.5, 2.6, 2.7, 2.8 and 2.9, both the radial position and velocity errors of Object ① present relatively high levels of divergence, and the standard deviation reached 35 km in the next day due to its relatively low orbital altitude and large influence of atmospheric damping.

Object ② presents a relatively mild degree of error divergence due to its relatively high orbital altitude and relatively small influence of atmospheric damping. In the next day, its standard deviations of position and velocity errors reached 27 km and 3.7 m/s, respectively.

As Object ③ is in an elliptic orbit, it is subject to the most complicated perturbation factors and the most serious error divergence situation. In the next day, its standard deviations of position and velocity errors reached 70 km and 60 m/s respectively. Due to the relatively large orbital eccentricity and flight-path angle, a relatively obvious difference in RSW and NTW coordinate systems has been found.

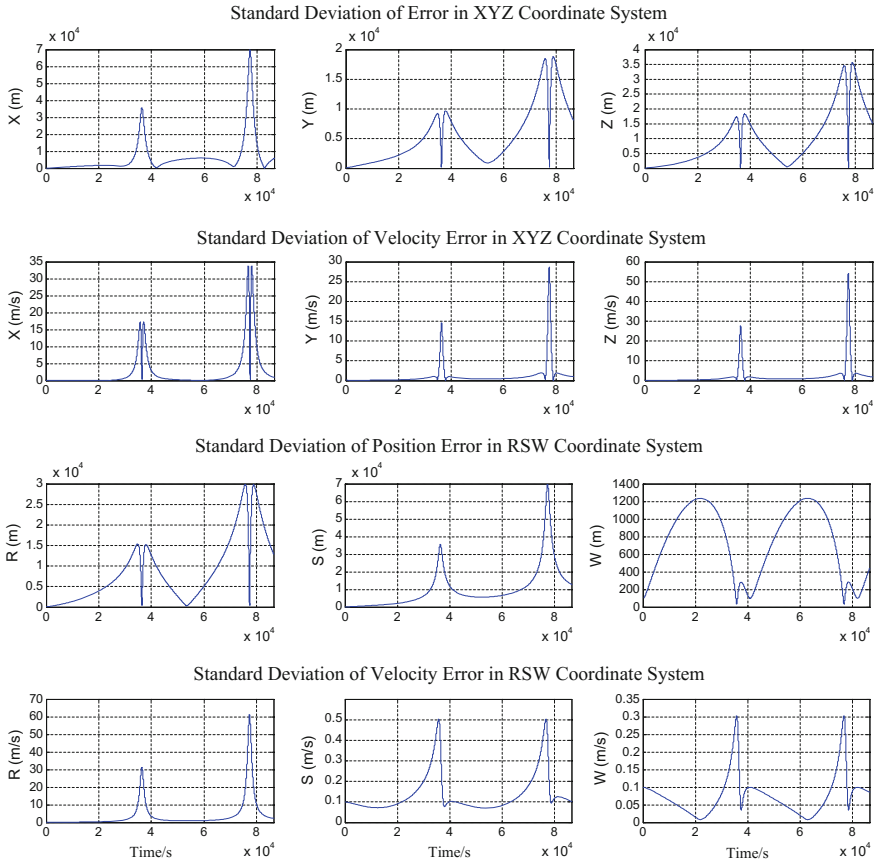


Fig. 2.7 Error propagation curve of object ③

The propagation curve of error standard deviation in NTW coordinate system had also been drawn. As can be seen in Fig. 2.10, NTW coordinate system is more applicable to elliptic orbits than RSW coordinate system.

Object ④ is with the highest orbital altitude and the smallest influence of atmospheric damping. Therefore it has the minimum error divergence in radial velocity, whose standard deviation only reached 2 m/s in the next day. However, it has a relatively serious error divergence in position, whose standard deviation reached 27 km in the next day.

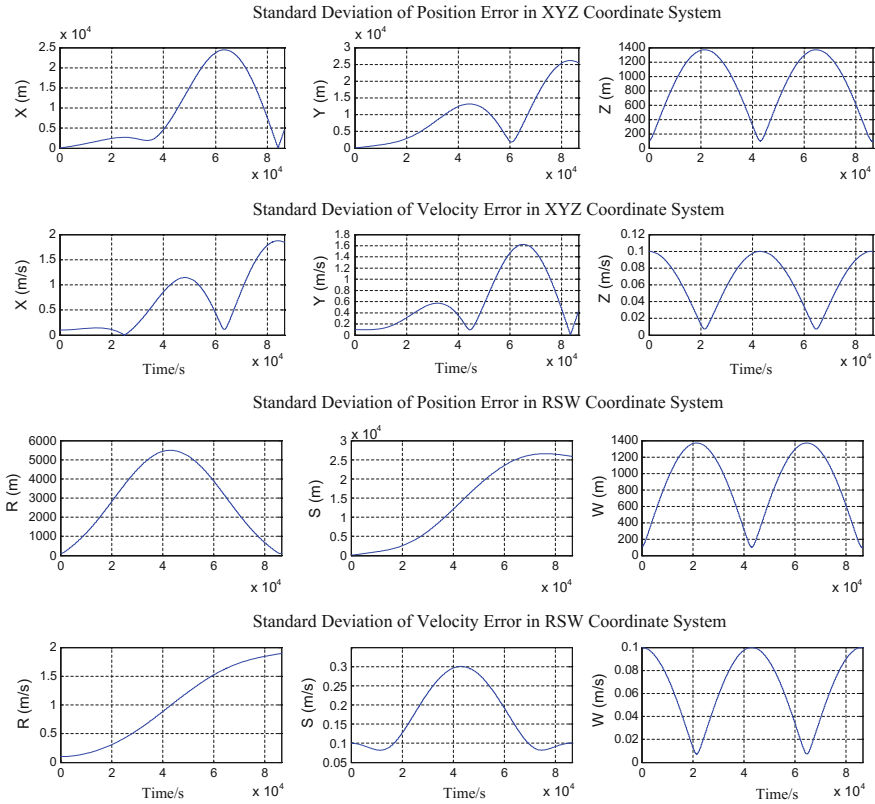


Fig. 2.8 Error propagation curve of object ④ (one day later)

2.4 Orbital Error Analysis Based on Relative Motion Theory

Relative motion of spacecraft refers to the motion of one or one set of spacecrafts relative to a target spacecraft or a virtual spacecraft under the action of the same gravitational field [21]. In contrast with its absolute motion (classical orbital motion), the relative motion of a spacecraft discusses questions in more variable forms and influenced by more complicated factors. Modeling and analysis of the relative motion of spacecraft had already gone through in-depth studies by various forms of models, which have been successfully applied to space missions including space rendezvous and docking, spacecraft in-orbit services, spacecraft formation flying, space-based space object surveillance, short distance observation, space robot capture, and so on [21].

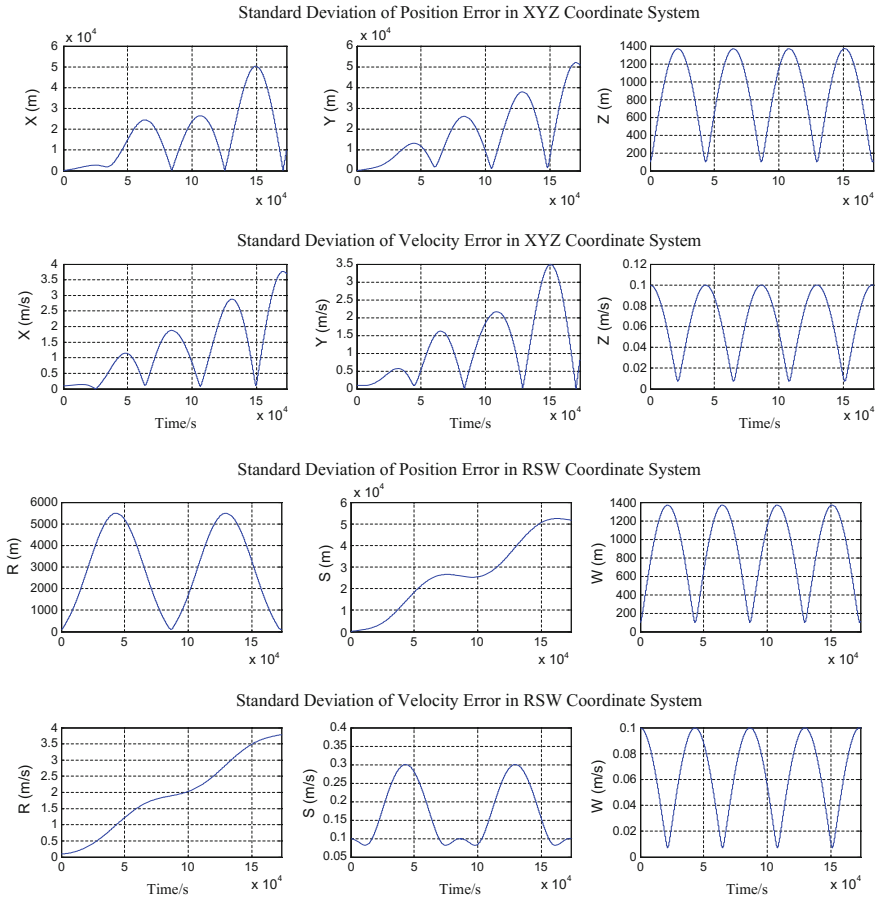


Fig. 2.9 Error propagation curve of object ④ (two days later)

Error (less than dozens of kilometers) of space object orbit propagation is a smaller value comparing to its radius vector (generally no less than 6500 km). Thus, the true and predict states of a space object can be considered as two objects with a short distance: a real and a virtual one, orbit prediction errors as relative motion between these two objects, then the relative motion theory in short distance can be used to study the characteristics of orbit prediction errors.

In this section, we firstly do initial error propagation analysis on circular and elliptical orbits using C-W and T-H equations based on algebraic models for relative motion. Then we make correlation characterization analysis on position and velocity prediction errors of nearly circular orbit using geometrical models for relative motion and then discuss its application in initial error selection.

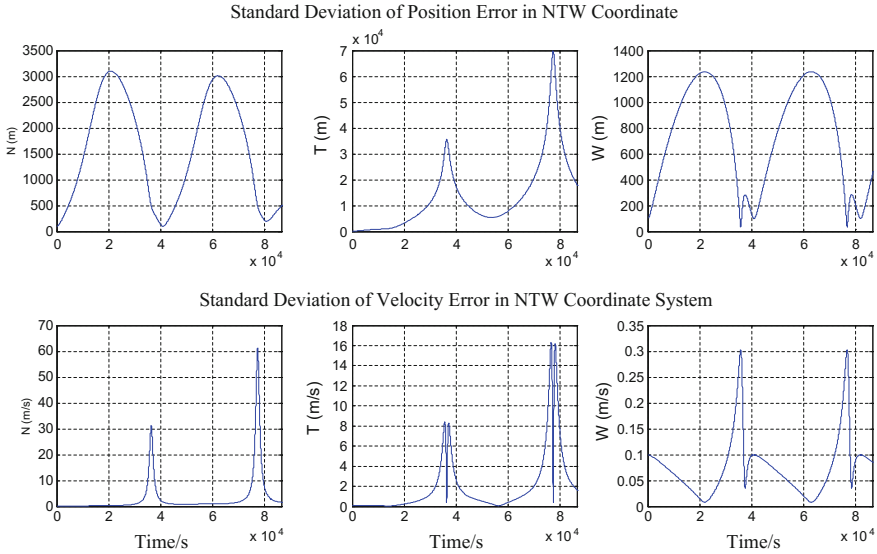


Fig. 2.10 Error propagation curve of object ③ in NTW coordinate system

2.4.1 Initial Error Propagation Based on Algebraic Models

Orbit prediction errors of space objects consist of initial and model errors. Initial error in orbit prediction propagates with time. Its propagation characteristics and tendencies differ with the type of orbits. Decoupling analysis on initial and model errors of orbit prediction is often used to simplify the problem. For initial error analysis, orbit prediction models are considered to be accurate without any model errors, which indicate that the difference of perturbation forces between prediction models and actual orbits is considered to be zero. Thus we can get the propagation characteristics of initial errors.

Generally, studies on relative motion neglect attitude motions. There are two types of models [22–26]. One is algebraic method, also known as the dynamic method, which is based on relative dynamic equation and easy for the guidance and control of relative motion. Another is geometrical method, also known as the kinematic method, which is based on models of relative motion expressed by orbital elements and easy for orbital design and perturbation analysis. The algebraic method is based on the basic dynamic equations described by absolute position vectors of two spacecraft and builds a relative motion model in relative motion coordinate system by assumptions and simplifications. The model is expressed in differential equation sets, and makes predictions for the relative motion state if initial conditions are known, so that it is applicable for analyzing the propagation characterization of initial errors of the orbit prediction.

1. Algebraic models for relative motion

The application of relative motion theory in analysis of orbit prediction errors of space object is discussed in this section. One of the two space objects is the dominant object and the other is the subordinate object, corresponding to the target spacecraft and the tracing spacecraft in space rendezvous and docking.

The algebraic models for relative motion has been developed by Hill [27] and Clohessy and Wiltshire [28]. Vallado [2] has presented the development as well. We build relative motion models based on algebraic method, starting from the basic motion equations described by absolute position vectors of the two objects and establish the relative motion model in the *RSW* coordinate system of the dominating object. If we take \mathbf{r}_p and \mathbf{r}_s as the respective absolute position vectors of the dominating and subordinate objects in *ECI* coordinate system, the orbital dynamic equations of the dominating and subordinate objects in *ECI* coordinate system are as follows

$$\ddot{\mathbf{r}}_p = -\frac{\mu}{r_p^3}\mathbf{r}_p + \mathbf{f}_p, \quad \ddot{\mathbf{r}}_s = -\frac{\mu}{r_s^3}\mathbf{r}_s + \mathbf{f}_s \quad (2.55)$$

where μ is constant of earth gravitation; \mathbf{f}_p and \mathbf{f}_s are respectively perturbation accelerations generated by all perturbation forces excluding gravity of the center of the Earth acting on the dominating and subordinate objects. For error analysis, the two objects are considered to be without control forces acting on them.

Relative position vector from the dominating object to the subordinate object $\mathbf{r} = \mathbf{r}_s - \mathbf{r}_p$, from Eq. (2.55) we can get

$$\frac{d^2\mathbf{r}}{dt^2} = \ddot{\mathbf{r}}_s - \ddot{\mathbf{r}}_p = -\frac{\mu}{r_s^3}\mathbf{r}_s + \frac{\mu}{r_p^3}\mathbf{r}_p + \mathbf{f}_s - \mathbf{f}_p = \frac{\mu}{r_p^3}\left(\mathbf{r}_p - \left(\frac{r_p}{r_s}\right)^3\mathbf{r}_s\right) + \mathbf{f} \quad (2.56)$$

where $\mathbf{f} = \mathbf{f}_s - \mathbf{f}_p$ is the difference of the perturbation accelerations acting on the two objects.

Equation (2.56) is the expression of relative dynamic equation in *ECI* coordinate system, which is an accurate equation of relative motion without any simplification.

Build a relative dynamic equation in *RSW* coordinate system of the dominating object. The relation of the absolute derivative and relative derivative is as follows:

$$\frac{d^2\mathbf{r}}{dt^2} = \frac{\delta^2\mathbf{r}}{\delta t^2} + 2\boldsymbol{\omega}_p \times \frac{\delta\mathbf{r}}{\delta t} + \boldsymbol{\omega}_p \times (\boldsymbol{\omega}_p \times \mathbf{r}) + \dot{\boldsymbol{\omega}}_p \times \mathbf{r} \quad (2.57)$$

where $\delta^2\mathbf{r}/\delta t^2$ and $\delta\mathbf{r}/\delta t$ are the relative acceleration vector and relative velocity vector of the subordinate object in *RSW* coordinate system of the dominating object, respectively; $\dot{\boldsymbol{\omega}}_p$ and $\boldsymbol{\omega}_p$ are the angular acceleration vector and angular velocity vector of the dominating object, respectively.

From Eqs. (2.56) and (2.57) we can get

$$\frac{\delta^2 \mathbf{r}}{\delta t^2} + 2\boldsymbol{\omega}_p \times \frac{\delta \mathbf{r}}{\delta t} + \boldsymbol{\omega}_p \times (\boldsymbol{\omega}_p \times \mathbf{r}) + \dot{\boldsymbol{\omega}}_p \times \mathbf{r} = \frac{\mu}{r_p^3} \left(\mathbf{r}_p - \left(\frac{r_p}{r_s} \right)^3 \mathbf{r}_s \right) + \mathbf{f} \quad (2.58)$$

In *RSW* coordinate system of the dominating object, assume the relative position vector to be $\mathbf{r} = (x \ y \ z)^T_{RSW}$, we can get

$$\begin{aligned} \frac{\delta \mathbf{r}}{\delta t} &= (\dot{x} \ \dot{y} \ \dot{z})^T, \quad \frac{\delta^2 \mathbf{r}}{\delta t^2} = (\ddot{x} \ \ddot{y} \ \ddot{z})^T \\ \boldsymbol{\omega}_p &= (0 \ 0 \ \dot{\theta})^T, \quad \dot{\boldsymbol{\omega}}_p = (0 \ 0 \ \ddot{\theta})^T \\ \mathbf{r}_p &= (r_p \ 0 \ 0)^T, \quad \mathbf{r}_s = (r_p + x \ y \ z)^T \end{aligned} \quad (2.59)$$

where θ is the true anomaly of the dominating object orbit.

Substitute Eq. (2.59) into Eq. (2.58) yields

$$\begin{cases} \ddot{x} = 2\dot{y}\dot{\theta} + x\dot{\theta}^2 + y\ddot{\theta} + \frac{\mu}{r_p^2} - \frac{\mu}{r_s^3}(r_p + x) + f_x \\ \ddot{y} = -2\dot{x}\dot{\theta} + y\dot{\theta}^2 - x\ddot{\theta} - \frac{\mu}{r_s^3}y + f_y \\ \ddot{z} = -\frac{\mu}{r_s^3}z + f_z \end{cases} \quad (2.60)$$

Equation (2.60) is the expression of relative dynamic equation in *RSW* coordinate system, which is an accurate equation of relative motion without any simplification.

Geocentric distance of the subordinate object can be expressed with geocentric distance of the dominating object r_p and the relative position vector \mathbf{r}

$$r_s = \left[(r_p + x)^2 + y^2 + z^2 \right]^{\frac{1}{2}} = \left(r_p^2 + 2r_p x + r^2 \right)^{\frac{1}{2}} = r_p \left(1 + 2\frac{x}{r_p} + \frac{r^2}{r_p^2} \right)^{\frac{1}{2}} \quad (2.61)$$

Suppose that the distance between the dominating and subordinate objects is far less than their geocentric distance ($r \ll r_p$), Eq. (2.61) can be written as

$$\frac{1}{r_s^3} = \frac{1}{r_p^3} \left(1 + 2\frac{x}{r_p} + \frac{r^2}{r_p^2} \right)^{-\frac{3}{2}} \approx \frac{1}{r_p^3} \left(1 - \frac{3x}{r_p} \right) \quad (2.62)$$

Substitute Eq. (2.62) into Eq. (2.60) yields

$$\begin{cases} \ddot{x} = 2\dot{y}\dot{\theta} + x\dot{\theta}^2 + y\ddot{\theta} + \frac{2\mu x}{r_p^3} + \frac{3\mu x^2}{r_p^4} + f_x \\ \ddot{y} = -2\dot{x}\dot{\theta} + y\dot{\theta}^2 - x\ddot{\theta} - \frac{\mu y}{r_p^3} + \frac{3\mu xy}{r_p^4} + f_y \\ \ddot{z} = -\frac{\mu z}{r_p^3} + \frac{3\mu xz}{r_p^4} + f_z \end{cases} \quad (2.63)$$

Furthermore, neglecting the quadratic term ($\mu x r_p^4$) in Eq. (2.63), we can get a linearized expression of relative motion equation

$$\begin{cases} \ddot{x} - 2\dot{\theta}\dot{y} - \dot{\theta}^2 x - \ddot{\theta}y - \frac{2\mu x}{r_p^3} = f_x \\ \ddot{y} + 2\dot{\theta}\dot{x} + \ddot{\theta}x - \dot{\theta}^2 y + \frac{\mu y}{r_p^3} = f_y \\ \ddot{z} + \frac{\mu z}{r_p^3} = f_z \end{cases} \quad (2.64)$$

Deduction of Eq. (2.64) only needs to suppose that the distance between the dominating and subordinate objects is far less than their geocentric distances, no restrictions on orbital eccentricity is required, and it is applicable to all orbital shapes. From Eq. (2.64) we can tell that the relative motion along the direction W (vertical to the orbital plane) is independent whereas the relative motions along the direction R and S are coupled with each other.

(1) C-W equation

If the dominating object is in a nearly circular orbit, i.e. $e_p \approx 0$, we have

$$\dot{\theta} \approx n, \quad \ddot{\theta} \approx 0, \quad \frac{\mu}{r_p^3} \approx n^2 \quad (2.65)$$

where n is average angular velocity of the dominating object.

Substituting Eq. (2.65) into Eq. (2.64) and neglecting the difference of perturbation forces of the two objects, we can get

$$\begin{cases} \ddot{x} - 2n\dot{y} - 3n^2 x = 0 \\ \ddot{y} + 2n\dot{x} = 0 \\ \ddot{z} + n^2 z = 0 \end{cases} \quad (2.66)$$

Equation (2.66) is the Clohessy-Wiltshire or C-W equation, also known as Hill's equation or H-C-W equation [2].

If we mark the initial time of relative motion with subscript "0", the corresponding initial conditions is $(x_0, y_0, z_0, \dot{x}_0, \dot{y}_0, \dot{z}_0)$ and we can get the analytical solution of relative motion at any time t by integrating Eq. (2.66). The analytical solution of relative position is

$$\begin{cases} x(t) = \frac{\dot{x}_0}{n} \sin nt - \left(\frac{2\dot{y}_0}{n} + 3x_0\right) \cos nt + 2\left(\frac{\dot{y}_0}{n} + 2x_0\right) \\ y(t) = 2\left(\frac{2\dot{y}_0}{n} + 3x_0\right) \sin nt + \frac{2\dot{x}_0}{n} \cos nt - 3(\dot{y}_0 + 2nx_0)t + \left(y_0 - \frac{2\dot{x}_0}{n}\right) \\ z(t) = \frac{\dot{z}_0}{n} \sin nt + z_0 \cos nt \end{cases} \quad (2.67)$$

The analytical solution of relative velocity is

$$\begin{cases} \dot{x}(t) = (2\dot{y}_0 + 3nx_0) \sin nt + \dot{x}_0 \cos nt \\ \dot{y}(t) = -2\dot{x}_0 \sin nt + 2(2\dot{y}_0 + 3nx_0) \cos nt - 3(\dot{y}_0 + 2nx_0) \\ \dot{z}(t) = -nz_0 \sin nt + \dot{z}_0 \cos nt \end{cases} \quad (2.68)$$

C-W equation is the most classical algebraic relative dynamic model. It takes the spherical central gravitation and circular reference orbit as assumptions, excluding the influence of perturbation forces. Although bearing some errors, it provides the simplest relative motion relationship and is widely used in satellite formation flying, space rendezvous, and other fields.

(2) T-H equation

In order to eliminate the restriction of circular reference orbit, Tschauner and Hempel assumed that the object distance is a first order small value comparing to the geocentric distance. According to variable substitution, they made the true anomaly of the dominating object an independent variable, thus the equation became dimensionless. Next, they simplified the nonlinear relative dynamic differential equation set into a linear time-varying equation set and got an analytical solution described with true anomaly/eccentric anomaly. The equation is called the Tschauner-Hempel Equation. (abbr. T-H equation). T-H equation describes the relative motion of objects with any eccentricity. Yamanaka got a new state transition matrix through a series of variable substitutions [29], which helped in discussion of its consistency with C-W equations. Yue et al., deduced specific expressions of state transition matrix factors expressed by positions and velocities in three-dimensional space. Converting them from angular domain into time domain, they got state transition matrix expressing actual positions and velocities [30]. Xing et al., studied relative motion issues of satellite formation flying [31–37].

If the dominating object is in an elliptical orbit, we can get a relatively simple dynamics equation by taking true anomaly θ as an independent variable. Therefore, we can get relative dynamics equations with true anomaly θ as an independent variable by making variable substitutions to Eq. (2.64) and substituting true anomaly θ for time t as the independent variable, derivative x' with respect to true anomaly for derivative \dot{x} with respect to time

$$\begin{cases} x'' - \frac{2e \sin \theta}{1+e \cos \theta} x' - 2y' - \frac{3+e \cos \theta}{1+e \cos \theta} x + \frac{2e \sin \theta}{1+e \cos \theta} y = 0 \\ y'' + 2x' - \frac{2e \sin \theta}{1+e \cos \theta} y' - \frac{2e \sin \theta}{1+e \cos \theta} x - \frac{e \cos \theta}{1+e \cos \theta} y = 0 \\ z'' - \frac{2e \sin \theta}{1+e \cos \theta} z' + \frac{1}{1+e \cos \theta} z = 0 \end{cases} \quad (2.69)$$

Yamanaka made variable substitutions to Eq. (2.69) [29], with its equations given in LVLH coordinate system, whereas equations of this section given in RSW coordinate system. If we make $\rho = 1 + e \cos \theta$, variable substitutions are as follows:

$$(\tilde{x} \ \tilde{y} \ \tilde{z})^T = \rho(x \ y \ z)^T \quad (2.70)$$

Converting T-H equations to a simpler form

$$\begin{cases} \tilde{x}'' = 3\tilde{x}/\rho + 2\tilde{y}' \\ \tilde{y}'' = -2x' \\ \tilde{z}'' = -\tilde{z} \end{cases} \quad (2.71)$$

Yamanaka, by derivation, got a state transition matrix expressed in angular domain [29]. We have deduced the expression of state transition matrix of T-H equation in the *RSW* coordinate system:

$$\begin{bmatrix} x \\ y \\ z \\ \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \mathbf{T}(t) \begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ \dot{x}_0 \\ \dot{y}_0 \\ \dot{z}_0 \end{bmatrix}, \quad \mathbf{T}(t) = \begin{bmatrix} t_{11} & t_{12} & 0 & t_{14} & t_{15} & 0 \\ t_{21} & t_{22} & 0 & t_{24} & t_{25} & 0 \\ 0 & 0 & t_{33} & 0 & 0 & t_{36} \\ t_{41} & t_{42} & 0 & t_{44} & t_{45} & 0 \\ t_{51} & t_{52} & 0 & t_{54} & t_{55} & 0 \\ 0 & 0 & t_{63} & 0 & 0 & t_{66} \end{bmatrix} \quad (2.72)$$

Due to the length limitation of this section, we list specific element values of the state transition matrix $\mathbf{T}(t)$ in Appendix A. The only assumption that makes the state transition matrix tenable is that the relative distance is far less than the geocentric distances of space objects and hence it can be applicable to elliptical orbit with any eccentricity. When we have already known initial conditions, the matrix can be directly used to calculate the state transition of the relative motion in the time domain.

2. Error propagation of near-circular orbit based on C-W equation

(1) Error propagation formula

We first use C-W equation to analyze propagation characteristics of initial errors of nearly circular orbits. $(\dot{x}, \dot{y}, \dot{z})^T$ in Eq. (2.68) can be seen as relative velocity errors or time derivatives of relative position errors, projection of absolute velocity errors in the orbital coordination system, which is more concerned in error analysis. From the following relation of absolute and relative velocities,

$$\frac{d\mathbf{r}}{dt} = \frac{\delta\mathbf{r}}{\delta t} + \boldsymbol{\omega}_p \times \mathbf{r} \quad (2.73)$$

we can get

$$\Delta\mathbf{v}_{RSW} = \frac{\delta\mathbf{r}}{\delta t} + \boldsymbol{\omega}_p \times \mathbf{r} = \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ n \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \dot{x} - ny \\ \dot{y} + nx \\ \dot{z} \end{pmatrix} \quad (2.74)$$

Finally,

$$\begin{cases} \Delta v_x = -(2\dot{y}_0 + 3nx_0) \sin nt - \dot{x}_0 \cos nt + 3n(\dot{y}_0 + 2nx_0)t + (2\dot{x}_0 - ny_0) \\ \Delta v_y = -\dot{x}_0 \sin nt + (2\dot{y}_0 + 3nx_0) \cos nt - (\dot{y}_0 + 2nx_0) \\ \Delta v_z = -nz_0 \sin nt + \dot{z}_0 \cos nt \end{cases} \quad (2.75)$$

Equation (2.75), together with Eq. (2.67), gives analytical expressions of error changes between orbit prediction values and their truth values, providing that initial position and velocity errors are already known.

If we combine the analytic expression (2.67) for position errors of C-W equation and the analytic expression (2.75) for its velocity errors together, we can write them into a state transition matrix. From Eq. (2.74) we can get

$$\dot{x} = \Delta v_x + ny, \dot{y} = \Delta v_y - nx \quad (2.76)$$

The relation between the state vector of errors $(x, y, z, \dot{x}, \dot{y}, \dot{z})^T$ and the state vector of errors $(x, y, z, \Delta v_x, \Delta v_y, \Delta v_z)^T$ is

$$\begin{bmatrix} x \\ y \\ z \\ \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & n & 0 & 1 & 0 & 0 \\ -n & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ \Delta v_x \\ \Delta v_y \\ \Delta v_z \end{bmatrix} = \mathbf{M}_{\text{abs} \rightarrow \text{rel}}(n) \begin{bmatrix} x \\ y \\ z \\ \Delta v_x \\ \Delta v_y \\ \Delta v_z \end{bmatrix} \quad (2.77)$$

Thus, the state transition matrix of state vector of error $\mathbf{X}(t) = (x, y, z, \Delta v_x, \Delta v_y, \Delta v_z)^T$ is

$$\Phi(t) = \begin{bmatrix} 2 - \cos nt & \sin nt & 0 & \frac{1}{n} \sin nt & \frac{2}{n} (1 - \cos nt) & 0 \\ 2 \sin nt - 3nt & 2 \cos nt - 1 & 0 & \frac{2}{n} (\cos nt - 1) & \frac{4}{n} \sin nt - 3t & 0 \\ 0 & 0 & \cos nt & 0 & 0 & \frac{1}{n} \sin nt \\ n(3nt - \sin nt) & n(1 - \cos nt) & 0 & 2 - \cos nt & 3nt - 2 \sin nt & 0 \\ n(\cos nt - 1) & -n \sin nt & 0 & -\sin nt & 2 \cos nt - 1 & 0 \\ 0 & 0 & -n \sin nt & 0 & 0 & \cos nt \end{bmatrix} \quad (2.78)$$

The expression of state transition matrix $\Phi(t)$ given in reference [38] is inaccurate, for it is a transition matrix from the initial state vector $(x_0, y_0, z_0, \dot{x}_0, \dot{y}_0, \dot{z}_0)^T$ to the terminal state vector $(x, y, z, \Delta v_x, \Delta v_y, \Delta v_z)^T$. $\Phi(t)$, multiplied by the matrix in Eq. (2.77) to its right, can get the transition matrix from $(x_0, y_0, z_0, \Delta v_{x0}, \Delta v_{y0}, \Delta v_{z0})^T$ to its terminal state $(x, y, z, \Delta v_x, \Delta v_y, \Delta v_z)^T$.

If we have already known the error vector of initial state X_0 , the state vector of error at any time t can be solved using the following equation

$$X(t) = \Phi(t)X_0 \quad (2.79)$$

Suppose that P_0 is the error covariance matrix in RSW coordinate system at the initial time t_0 , then the covariance matrix at any time t is

$$P(t) = \Phi(t) \cdot P_0 \cdot \Phi^T(t) \quad (2.80)$$

If initial covariance matrix P_0 is diagonal matrix,

$$P_0 = \text{diag}\left(\sigma_{x_0}^2 \quad \sigma_{y_0}^2 \quad \sigma_{z_0}^2 \quad \sigma_{v_x,0}^2 \quad \sigma_{v_y,0}^2 \quad \sigma_{v_z,0}^2\right) \quad (2.81)$$

Substitute it into Eq. (2.80) and use Eq. (2.78), we can solve the covariance matrix at any time t

$$P(t) = \begin{bmatrix} P_{rr}(t) & P_{rv}(t) \\ P_{rv}^T(t) & P_{vv}(t) \end{bmatrix} \quad (2.82)$$

where, $P_{rr}(t)$, $P_{rv}(t)$, and $P_{vv}(t)$ stand for position covariance matrix, cross correlation matrix of position and velocity, and velocity covariance matrix, respectively. Refer to Appendix B for their specific expressions. In addition, Ref. [38] points out that the covariance matrix at any time t is a diagonal matrix, so that the conclusion is inaccurate.

The following is the expression of the diagonal element of $P(t)$:

$$\left\{ \begin{array}{l} \sigma_x^2(t) = (\cos nt - 2)^2 \sigma_{x_0}^2 + \sin^2 nt \sigma_{y_0}^2 + \frac{\sin^2 nt}{n^2} \sigma_{v_x,0}^2 + \frac{4(1 - \cos nt)^2}{n^2} \sigma_{v_y,0}^2 \\ \sigma_y^2(t) = (2 \sin nt - 3nt)^2 \sigma_{x_0}^2 + (2 \cos nt - 1)^2 \sigma_{y_0}^2 + \frac{4(1 - \cos nt)^2}{n^2} \sigma_{v_x,0}^2 \\ \quad + \left(\frac{4}{n} \sin nt - 3t\right)^2 \sigma_{v_y,0}^2 \\ \sigma_z^2(t) = \cos^2 nt \sigma_{z_0}^2 + \frac{1}{n^2} \sin^2 nt \sigma_{v_z,0}^2 \\ \sigma_{v_x}^2(t) = n^2 (\sin nt - 3nt)^2 \sigma_{x_0}^2 + n^2 (\cos nt - 1)^2 \sigma_{y_0}^2 + (2 - \cos nt)^2 \sigma_{v_x,0}^2 \\ \quad + (3nt - 2 \sin nt)^2 \sigma_{v_y,0}^2 \\ \sigma_{v_y}^2(t) = n^2 (1 - \cos nt)^2 \sigma_{x_0}^2 + n^2 \sin^2 nt \sigma_{y_0}^2 + \sin^2 nt \sigma_{v_x,0}^2 \\ \quad + (2 \cos nt - 1)^2 \sigma_{v_y,0}^2 \\ \sigma_{v_z}^2(t) = n^2 \sin^2 nt \sigma_{z_0}^2 + \cos^2 nt \sigma_{v_z,0}^2 \end{array} \right. \quad (2.83)$$

Thus, by C-W equations, we can get the analytic propagation function of the initial error covariance matrix. From Eq. (2.83) we can know that, under assumptions of C-W equations, the initial error propagation of nearly circular orbit possesses the following characteristics:

- (1) Periodic oscillations exist in errors at all directions (positions and velocities).
- (2) Errors along direction W are independent to those in the orbital plane (along direction R and S) so that the errors can be resolved in the orbital plane and along the normal direction, respectively;
- (3) Errors along direction W do not have any constant terms, only sine and cosine function terms, which define free oscillations of errors along direction W ;
- (4) Except for oscillation terms, position errors along direction $S(\sigma_y^2)$, velocity errors along direction $R(\sigma_{v_x}^2)$ and the polynomial term of time t explain that their errors will increase with time and cause error divergence.

(2) Example of verification

We choose four typical objects as calculation examples to verify the orbital initial error propagation analytic formula based on the relative motion theory. See Table 2.3 for type of objects and orbits. See Tables 2.4 and 2.5 for orbital elements of objects as well as position and velocity coordinates in ECI coordinate system (Coordinated Universal Time is 2012-01-01 00:00:00.00 UTC).

Table 2.3 Types of objects and orbits

No.	Type of orbit	Type of object	Object name	Object code
①	Low Earth orbit (LEO)	Earth Resources Satellite	ERS-2	23560
②	Medium Earth orbit (MEO)	GLONASS Satellite	COSMOS 1948	19165
③	Geostationary Earth orbit (GEO)	Galaxy Communication Satellite	Galaxy-11	26038
④	Highly elliptical orbit (HEO)	Molnyia Satellite	Molnyia 3-35	20052

Table 2.4 Orbital elements of objects

No.	a/km	e	$i/(^\circ)$	$\Omega / (^\circ)$	$\omega / (^\circ)$	$M/(^\circ)$
①	7151.799943	0.002586	98.618	77.249	76.679	177.919
②	25507.551932	0.000375	64.354	242.658	306.373	53.654
③	42166.577215	0.000036	0.068	109.622	217.823	41.434
④	25663.740026	0.740361	63.463	359.792	270.021	40.168

Table 2.5 Position and velocity coordinates in ECI coordinate system

No.	X/km	Y/km	Z/km	V_x /(km/s)	V_y /(km/s)	V_z /(km/s)
①	-1430.331391	-1627.484800	-6835.083611	1.295564	7.067381	-1.954642
②	-11702.734586	-22658.145307	24.604748	1.521520	-0.783326	3.564419
③	41659.812155	6510.137799	-49.166290	-0.474640	3.037799	-0.000680
④	17110.947418	6872.467973	13886.072215	0.405454	1.938054	3.883779

Suppose that the standard deviation of position errors along each axis directions in *RSW* coordinate system at an initial time is 100 m and the standard deviation of velocity errors is 0.1 m/s, the initial covariance matrix will be

$$\mathbf{P}_0 = \text{diag}(10^4 \quad 10^4 \quad 10^4 \quad 10^{-2} \quad 10^{-2} \quad 10^{-2}) \quad (2.84)$$

In order to verify the correctness of Eq. (2.82), we use the following three methods to analyze the initial error propagation:

- (1) The HPOP model in *STK* software, 21×21 order WGS84_EGM96 model of earth gravitation, air drag coefficient 2.2, surface-mass ratio $0.02 \text{ m}^2/\text{kg}$, Jacchia-Roberts model of atmospheric density, solar radiation pressure coefficient 1.0. We also choose 21×21 order of gravitational field for the covariance calculation in HPOP model. This method only gives position error covariance, hereinafter referred to as HPOP.
- (2) The Monte-Carlo method considering only the two-body problem of which the number of sample points is $n_{\text{M-C}} = 10^4$, hereinafter referred to as M-C.
- (3) Error propagation analytic formula based on C-W equation, hereinafter referred to as C-W.

The time steps in three methods are all 1 min. It is necessary to illustrate that there are two approaches to generate initial random state with Monte-Carlo method. In the first approach, we generate random sample points at time t_0 which follow the distribution of initial errors then predict these sample points to each time t_i ($i = 1, 2, \dots$) using a two-body orbital model. We calculate the covariance of sample prediction values at each time t_i to get information of errors at that time. In the other approach, we regenerate $n_{\text{M-C}}$ initial sample points at time t_0 when we calculate sample prediction values at time t_i . Results of both approaches are basically the same, while the first one gets a smoother standard deviation propagation curve. If we consider from the aspect of analyzing initial error propagation, the first one can explain the changes of initial errors at each time better. Thus we choose the first approach to perform the Monte-Carlo statistical test.

See Figs. 2.11, 2.12 and 2.13 for error standard deviation divergence of three objects (objects ①–③) on nearly circular orbits in *RSW* coordinate system. We can learn that position errors along direction *S* and velocity errors along direction *R* on nearly circular orbit have divergences as time passes. Position and velocity errors along other directions have periodic oscillations as time passes without any divergences.

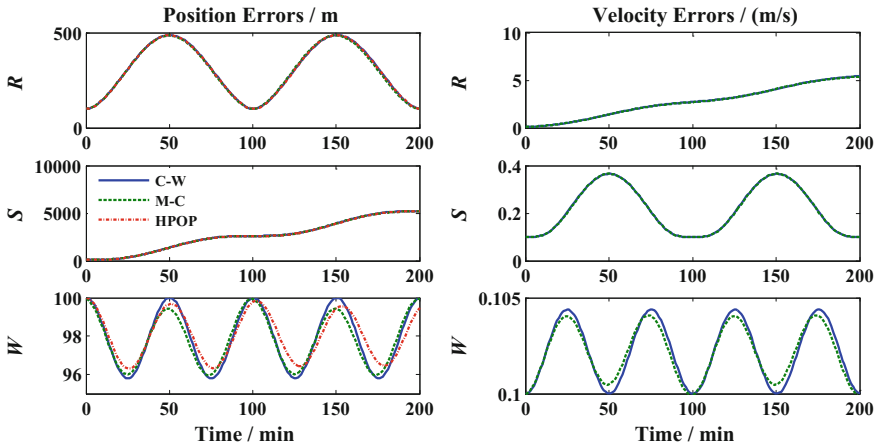


Fig. 2.11 Propagation curve of error standard deviations of object ①

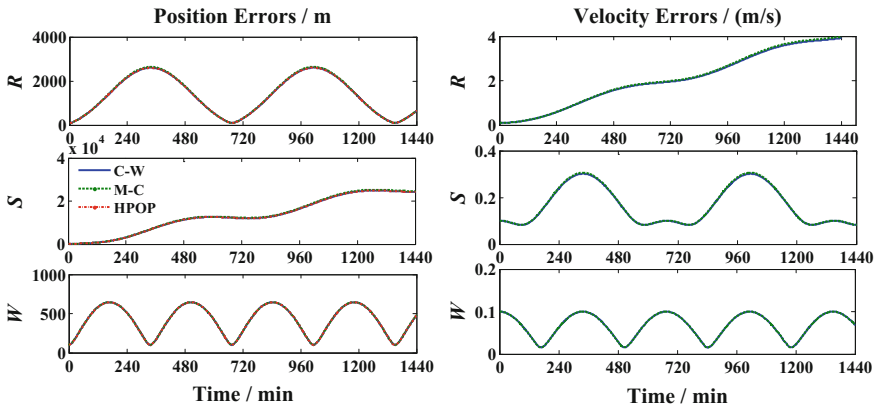


Fig. 2.12 Propagation curve of error standard deviations of object ②

Error standard deviation curves we get from the three methods in Figs. 2.11, 2.12 and 2.13 are very close and hard to distinguish. In order to analyze precisions with C-W method, we list relative error curves with C-W, M-C and HPOP methods in Figs. 2.14 2.15 and 2.16.

We can learn from the derivational process of C-W equation that there are three approximate hypotheses if we describe relative motion of nearly circular orbits with C-W Eq. (2.66). The first one is perturbation force hypothesis, in which the two objects are considered with the same perturbation force and their difference is zero. The second one is linear hypothesis, in which the relative distance of the two objects is considered as a smaller value comparing to their geocentric distance and the equation can be linearized. The third one is the circular orbit hypothesis, in

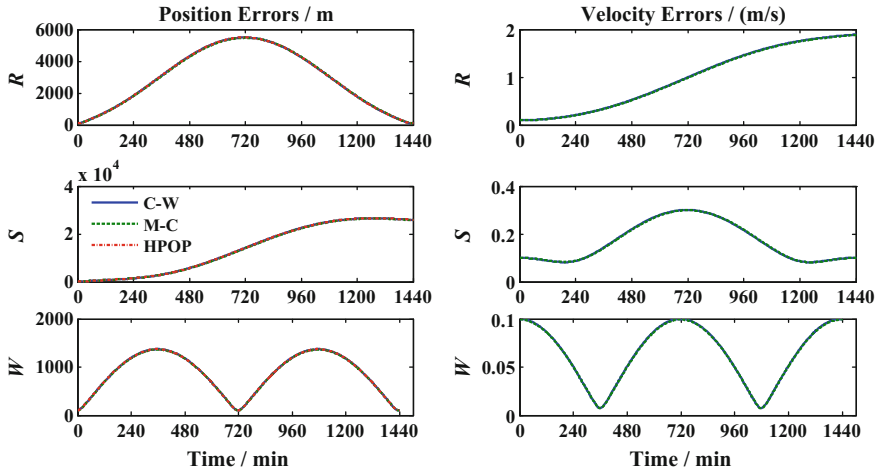


Fig. 2.13 Propagation curve of error standard deviations of object ③

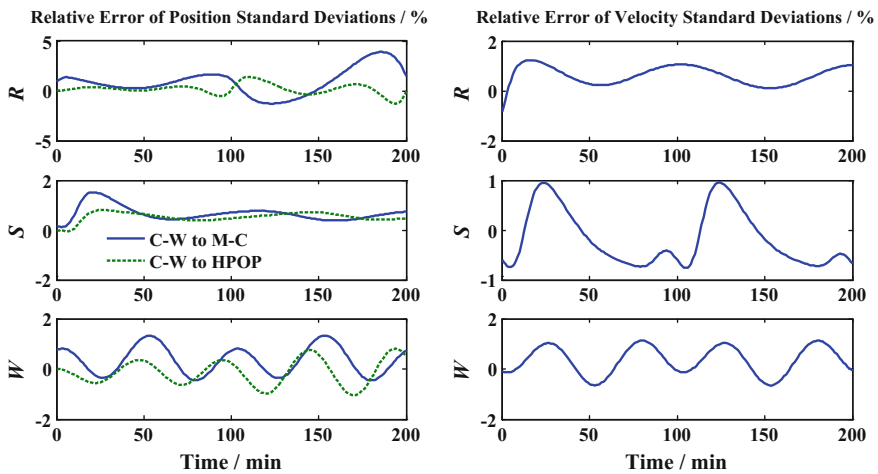


Fig. 2.14 Relative error curve of standard deviations of object ①

which the dominating object orbit is considered as strict circle and the equation can be converted into constant coefficient linear equation.

Likewise, analysis on nearly circular orbital error propagation with C-W equation is also affected by these three hypotheses. In Figs. 2.11, 2.12 and 2.13, deviations among C-W curves and M-C curves reveal deviations caused by linear and circular orbit hypotheses, while deviations among C-W curves and HPOP curves reveal deviations caused by the perturbation force hypothesis. If we analyze the comparison of results in Figs. 2.14, 2.15 and 2.16, we can find that the relative

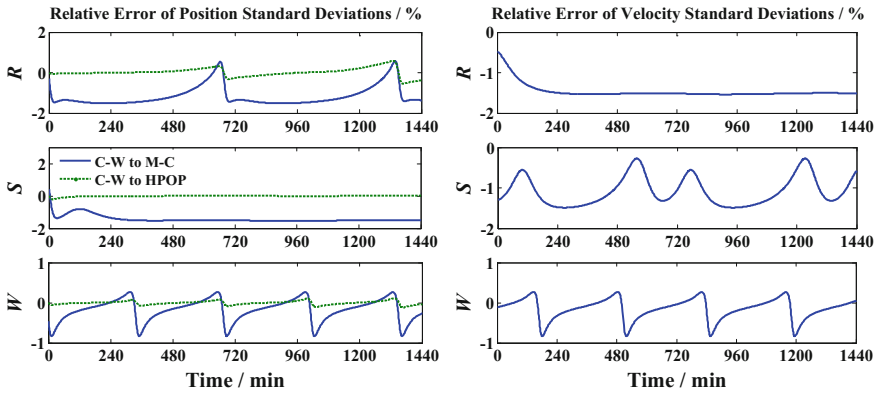


Fig. 2.15 Relative error curve of standard deviations of object ②

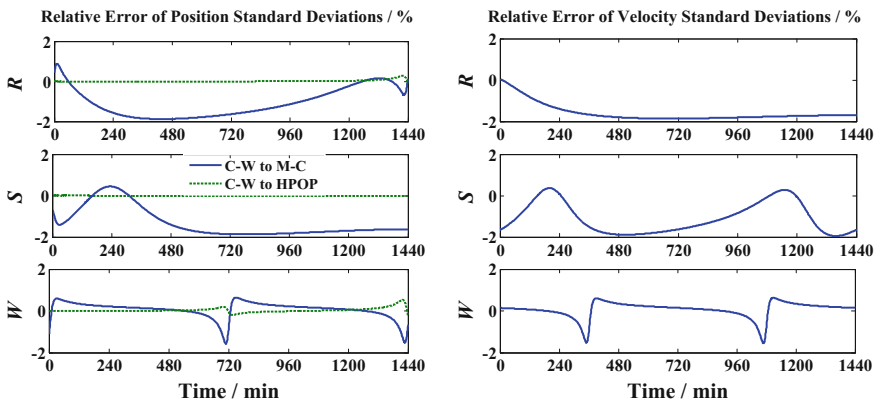


Fig. 2.16 Relative error curve of standard deviations of object ③

errors of results with C-W method are usually no larger than 2% comparing with those with M-C and HPOP methods, which indicates that the precision of analyzing initial errors with C-W equation is adequate.

From Fig. 2.11 we can learn that deviation of error standard deviation of object ① along direction W is relatively larger, which is due to its relatively smaller amplitude of variation. From Fig. 2.14 we can see that relative errors of object ① along direction W are smaller than 2%, no larger than that of other directions.

The relative errors are very small using C-M method and HPOP method in Figs. 2.14, 2.15 and 2.16. We can learn from the derivational process of C-W equation that two-body assumption is not introduced. The assumption is that perturbation forces are the same acting upon the dominating and subordinate objects and that the difference of perturbation forces is zero. When applied to error analysis, it can be seen as that the models of “actual” and prediction orbits are the same and

without any errors. Therefore, on the premise that C-W equation assumption has been met, what orbital model we choose has little to do with the results, which also explains that C-W equation is a good tool for initial error propagation analysis.

3. Error propagation of elliptical orbit based on T-H equation

(1) Error propagation formula

Likewise, the velocity component of prediction errors we get from Eq. (2.72) is relative velocity errors, or it can be seen as time derivative of relative position errors, whereas the projection of absolute velocity errors in the reference orbital coordinate system needs to be concerned. If we take $\Phi(t)$ for the state transition matrix of initial error propagation after its conversion to absolute velocity errors, it can be written as follows according to Eqs. (2.72), (2.74), and (2.76).

$$\begin{aligned} \Phi(t) &= \mathbf{M}_{\text{rel} \rightarrow \text{abs}}(\omega) \mathbf{T}(t) \mathbf{M}_{\text{abs} \rightarrow \text{rel}}(\omega_0) \\ &= \begin{bmatrix} t_{11} - \omega_0 t_{15} & t_{12} + \omega_0 t_{14} & 0 & t_{14} & t_{15} & 0 \\ t_{21} - \omega_0 t_{25} & t_{22} + \omega_0 t_{24} & 0 & t_{24} & t_{25} & 0 \\ 0 & 0 & t_{33} & 0 & 0 & t_{36} \\ t_{41} - \omega t_{21} - \omega_0(t_{45} - \omega t_{25}) & t_{42} - \omega t_{22} + \omega_0(t_{44} - \omega t_{24}) & 0 & t_{44} - \omega t_{24} & t_{45} - \omega t_{25} & 0 \\ t_{51} + \omega t_{11} - \omega_0(t_{55} + \omega t_{15}) & t_{52} + \omega t_{12} + \omega_0(t_{54} + \omega t_{14}) & 0 & t_{54} + \omega t_{14} & t_{55} + \omega t_{15} & 0 \\ 0 & 0 & t_{63} & 0 & 0 & t_{66} \end{bmatrix} \end{aligned} \quad (2.85)$$

If initial covariance matrix \mathbf{P}_0 is a diagonal matrix as in Eq. (2.81), we can get covariance matrix $\mathbf{P}(t) = \Phi(t) \mathbf{P}_0 \Phi^T(t)$ at any time t . Refer to Appendix B for the specific expression of covariance matrix $\mathbf{P}(t)$. Here we have the expression of its diagonal element.

$$\left\{ \begin{aligned} \sigma_x^2(t) &= (t_{11} - \omega_0 t_{15})^2 \sigma_{x0}^2 + (t_{12} + \omega_0 t_{14})^2 \sigma_{y0}^2 + t_{14}^2 \sigma_{v_x0}^2 + t_{15}^2 \sigma_{v_y0}^2 \\ \sigma_y^2(t) &= (t_{21} - \omega_0 t_{25})^2 \sigma_{x0}^2 + (t_{22} + \omega_0 t_{24})^2 \sigma_{y0}^2 + t_{24}^2 \sigma_{v_x0}^2 + t_{25}^2 \sigma_{v_y0}^2 \\ \sigma_z^2(t) &= t_{33}^2 \sigma_{z0}^2 + t_{36}^2 \sigma_{v_z0}^2 \\ \sigma_{v_x}^2(t) &= [t_{41} - \omega t_{21} - \omega_0(t_{45} - \omega t_{25})]^2 \sigma_{x0}^2 + (t_{44} - \omega t_{24})^2 \sigma_{v_x0}^2 \\ &\quad + [t_{42} - \omega t_{22} + \omega_0(t_{44} - \omega t_{24})]^2 \sigma_{y0}^2 + (t_{45} - \omega t_{25})^2 \sigma_{v_y0}^2 \\ \sigma_{v_y}^2(t) &= [t_{51} + \omega t_{11} - \omega_0(t_{55} + \omega t_{15})]^2 \sigma_{x0}^2 + (t_{54} + \omega t_{14})^2 \sigma_{v_x0}^2 \\ &\quad + [t_{52} + \omega t_{12} + \omega_0(t_{54} + \omega t_{14})]^2 \sigma_{y0}^2 + (t_{55} + \omega t_{15})^2 \sigma_{v_y0}^2 \\ \sigma_{v_z}^2(t) &= t_{63}^2 \sigma_{z0}^2 + t_{66}^2 \sigma_{v_z0}^2 \end{aligned} \right. \quad (2.86)$$

In which, t_{ij} is defined by orbital eccentricity e , true anomaly θ , and other elements. Refer to Appendix A for its specific expression. So we can get the explicit propagation functions of initial error covariance by the state transition matrix of elliptical orbits.

If we analyze propagation characteristics of initial errors, we should consider whether there are any divergences with time. Therefore, we divide propagation functions of initial error covariance into long-term items and periodic items for separate observations. The long-term items will cause error divergences whereas the periodic items will only make error oscillations within a certain scope. In order to see clearly the impacts of long-term and short items on initial error propagation, we rewrite coefficients of each item in Eq. (2.86) into the following:

$$t_{ij} = a_{ij}J + \Gamma_{ij}, \quad i, j = 1, 2, \dots, 6 \quad (2.87)$$

where $J = k^2(t - t_0)$ is a long-term item changing with time; a_{ij} is the coefficient of the long-term items; Γ_{ij} is the coefficient of the periodic items.

Therefore, we can rewrite the explicit propagation functions of initial error covariance into

$$\left\{ \begin{array}{l} \sigma_x^2(t) = [a_{11}J + \Gamma_{11} - \omega_0(a_{15}J + \Gamma_{15})]^2 \sigma_{x0}^2 + [a_{12}J + \Gamma_{12} + \omega_0(a_{14}J + \Gamma_{14})]^2 \sigma_{y0}^2 \\ \quad + (a_{14}J + \Gamma_{14})^2 \sigma_{v_x0}^2 + (a_{15}J + \Gamma_{15})^2 \sigma_{v_y0}^2 \\ \sigma_y^2(t) = [a_{21}J + \Gamma_{21} - \omega_0(a_{25}J + \Gamma_{25})]^2 \sigma_{x0}^2 + [a_{22}J + \Gamma_{22} + \omega_0(a_{24}J + \Gamma_{24})]^2 \sigma_{y0}^2 \\ \quad + (a_{24}J + \Gamma_{24})^2 \sigma_{v_x0}^2 + (a_{25}J + \Gamma_{25})^2 \sigma_{v_y0}^2 \\ \sigma_z^2(t) = \Gamma_{33}^2 \sigma_{z0}^2 + \Gamma_{36}^2 \sigma_{v_z0}^2 \\ \sigma_{v_x}^2(t) = [a_{41}J + \Gamma_{41} - \omega(a_{21}J + \Gamma_{21}) - \omega_0(a_{45}J + \Gamma_{45} - \omega(a_{25}J + \Gamma_{25}))]^2 \sigma_{x0}^2 \\ \quad + (a_{44}J + \Gamma_{44} - \omega(a_{24}J + \Gamma_{24}))^2 \sigma_{v_x0}^2 \\ \quad + [a_{42}J + \Gamma_{42} - \omega(a_{22}J + \Gamma_{22}) + \omega_0(a_{44}J + \Gamma_{44} - \omega(a_{24}J + \Gamma_{24}))]^2 \sigma_{y0}^2 \\ \quad + (a_{45}J + \Gamma_{45} - \omega(a_{25}J + \Gamma_{25}))^2 \sigma_{v_y0}^2 \\ \sigma_{v_y}^2(t) = [\Gamma_{51} + \omega \Gamma_{11} - \omega_0(\Gamma_{55} + \omega \Gamma_{15})]^2 \sigma_{x0}^2 + (\Gamma_{54} + \omega \Gamma_{14})^2 \sigma_{v_x0}^2 \\ \quad + [\Gamma_{52} + \omega \Gamma_{12} + \omega_0(\Gamma_{54} + \omega \Gamma_{14})]^2 \sigma_{y0}^2 + (\Gamma_{55} + \omega \Gamma_{15})^2 \sigma_{v_y0}^2 \\ \sigma_{v_z}^2(t) = \Gamma_{63}^2 \sigma_{z0}^2 + \Gamma_{66}^2 \sigma_{v_z0}^2 \end{array} \right. \quad (2.88)$$

where $a_{5j} + \omega a_{1j} = 0$ ($j = 1, 2, 4, 5$). Taking $j = 1$ for an example, we can prove the following using element values in Appendix A

$$\left\{ \begin{array}{l} \omega = h/r_T^2 = (h/p^2)(1 + e \cos \theta)^2 = k^2(1 + e \cos \theta)^2 = k^2 \rho^2 \\ a_{11} = \frac{3e\rho_0 s(e^2 + 3\rho_0 - 1)}{\rho(e^2 - 1)} - \frac{3e^3 s s_0^2}{\rho \rho_0(e^2 - 1)} \\ a_{51} = \frac{3e^3 k^2 \rho s s_0^2}{\rho_0(e^2 - 1)} - \frac{3ek^2 \rho \rho_0 s(e^2 + 3\rho_0 - 1)}{e^2 - 1} \end{array} \right. \quad (2.89)$$

From Eq. (2.89) we can get

$$a_{51} + \omega a_{11} = 0 \quad (2.90)$$

We can prove other equations in a similar way.

We can get propagation characteristics of elliptical orbit prediction errors using initial error propagation functions in Eq. (2.88):

- (1) Position errors along direction R and S in orbits contain long-term item J , which will increase as time passes.
- (2) Position errors along direction W in orbits only contain periodic items, which will not have any divergences as time passes.
- (3) Velocity errors along direction S and W in orbits will not have any divergences as time passes, whereas velocity errors along direction R will have divergences as time passes because they contain long-term item J .

We can also see from element values of state transition matrix in Annex A that there is eccentricity e in the members of long-term item coefficient a_{1j} ($j = 1, 2, 4, 5$) of position errors along direction R . Therefore, for circular orbits ($e = 0$), the long-term item coefficient of position errors along direction R is 0. That is to say, position errors along direction R will not have any divergences with time. In fact, for nearly circular orbits, $e \approx 0$, $\omega \approx n$, Eq. (2.88) can degrade into error propagation function Eq. (2.83) which we get from C-W equation.

(2) Example of verification

We choose typical space objects on highly eccentric orbits (object ④ in Table 2.3) and refer to Tables 2.3, 2.4 and 2.5 for orbital parameters. If we suppose that the standard deviation of position errors along each direction in RSW coordinate system at the initial time is 100 m and the standard deviation of velocity errors is 0.1 m/s, the initial covariance matrix will be consistent with that in Eq. (2.84). We analyze initial error propagation separately using the STK/HPOP model, the Monte-Carlo method which only takes two-body motion into account, and the error propagation analytic formula based on T-H equation, respectively. The time step is 1 min and the number of sample points is $nM-C = 104$ in the Monte-Carlo method. See Fig. 2.17 for error standard deviation divergence of object ④ in RSW coordination system.

As Object ④ is in an elliptic orbit, it is subject to complicated perturbation factors and with serious error divergences which cause acute changes to its error propagation curve. Position errors along direction R and S and velocity errors along direction R in the orbit continue to increase with time. In the next day, their standard deviations reached at about 30 km, 70 km, and 60 m/s, respectively. Position and velocity errors along other directions caused by initial errors have periodic oscillations as time passes without any divergences.

Error standard deviation curves we get from the three methods in Fig. 2.17 are very close and hard to distinguish. In order to analyze relative errors, we list their curves using T-H, M-C, and HPOP methods in Fig. 2.18.

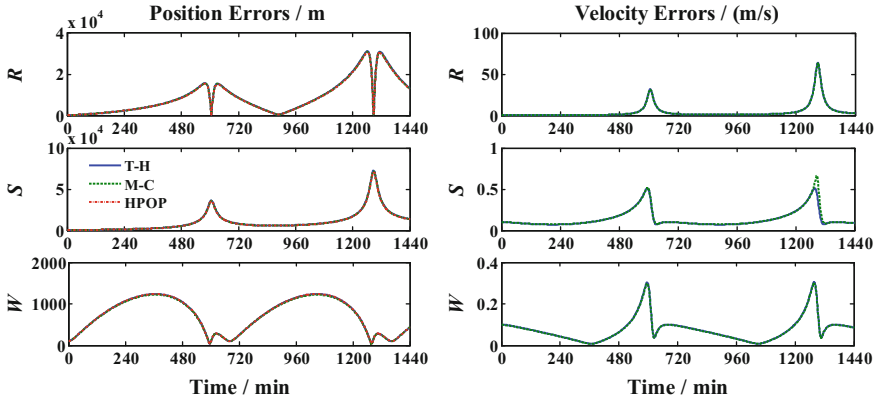


Fig. 2.17 Propagation curve of error standard deviations of object ④ (in RSW coordination system)

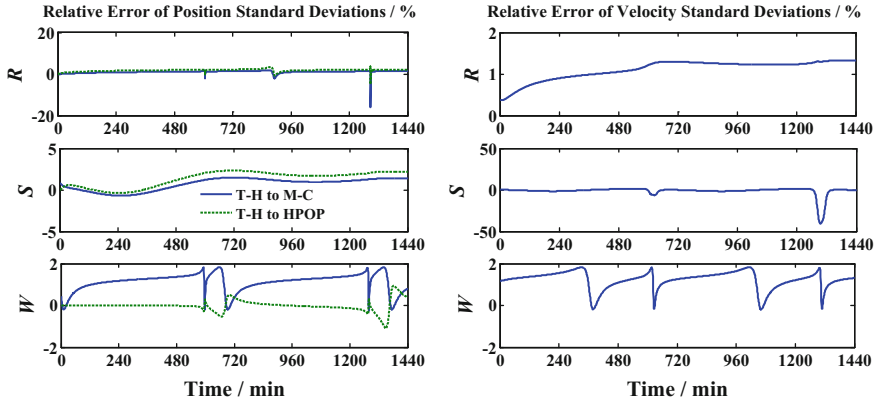


Fig. 2.18 Relative error curve of standard deviations of object ④ (in RSW coordination system)

From Figs. 2.17 and 2.18 we can learn that there is great difference between M-C and T-H curves at around 1290 s for standard deviations of velocity errors along direction S . The relative error reaches to 40% when objects are near the perigee. The possible reasons are that the curvature of elliptical orbits is relatively larger near the perigee and that sample points of velocity errors from Monte-Carlo method distribute in arc along orbits as in Fig. 2.19. The error distribution shows non-Gaussian features at that time whereas default errors in calculating error propagation using T-H method accord with Gaussian distribution. If we calculate covariance matrix directly, we may cause larger errors along direction S (i.e. the vertical axis in figures).

As in Fig. 2.20, $\mathbf{r} = \mathbf{r}_s - \mathbf{r}_p$ is error vector of the prediction position with respect to the actual position (or the relative position vector from the dominating object to

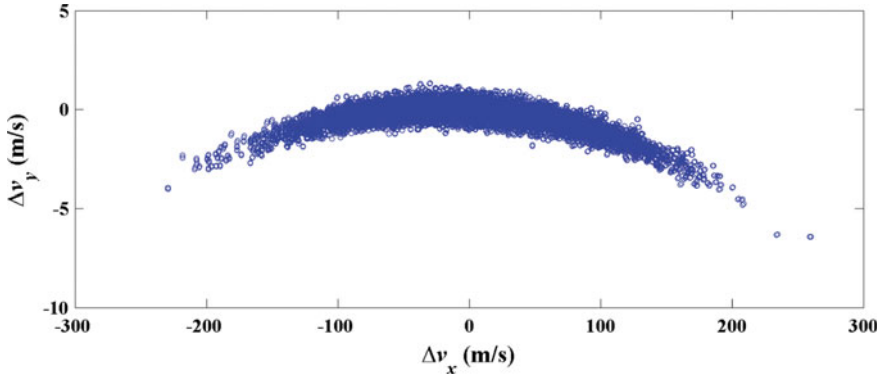


Fig. 2.19 Distribution of velocity errors using Monte-Carlo method at 1290s

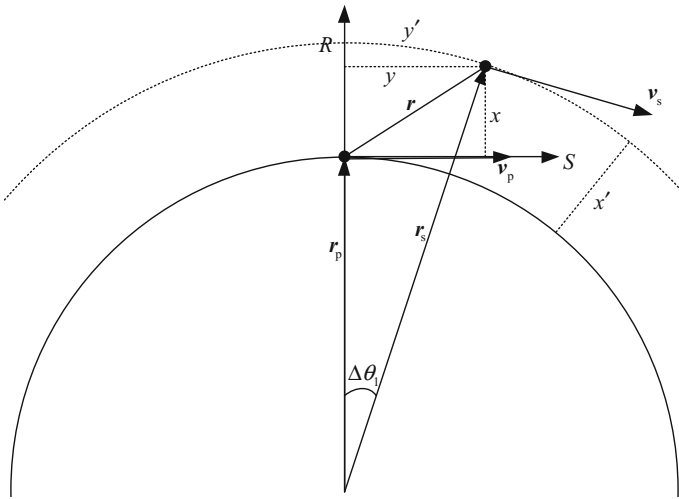


Fig. 2.20 Modifications of error data using orbit curvature

the subordinate object). We can describe error vector r as $r = [x, y, z]^T$ for real positions in RSW coordinate system. If we describe errors with the three components $[x, y, z]^T$ in RSW coordinate system (a rectangular coordinate system), we can get non-Gaussian distribution errors so that it will be improper to perform error fitting with Gaussian distribution assumption. However, when orbit curvature is not very large, the non-Gaussian features of orbital errors are not obvious. When orbit curvature is comparatively larger (as at the perigee of elliptical orbits), the non-Gaussian features of orbital errors are relatively obvious. We can modify error data using orbit curvature and describe it in a non-rectangular coordinate system.

As in Fig. 2.20, we define a polar coordinate system. The geocenter is taken as its origin and actual position r_p as its initial position. We can use an angle $\Delta\theta_1$ and a

distance $x' = r_s - r_p$ to present the two-dimensional error vector \mathbf{r} in the orbital plane, in which $\Delta\theta_1$ stands for a tiny angle between actual and prediction positions of the object. Modified results of error data are $\mathbf{r} = [x', y', z']^T$, where x' is the radial distance of prediction position away from its original orbit; y' is front or back deviations of prediction position in orbits; z' is the normal distance of prediction position away from its orbit plane. They can be respectively presented as follows

$$x' = r_s - r_p, \quad y' = r_p \Delta\theta_1, \quad z' = z \tag{2.91}$$

When orbit curvature is comparatively smaller, we can consider $x' = x$ and $y' = y$; when orbit curvature is comparatively larger, differences between x' and x , y' and y become non-negligible. Similarly, differences between velocity errors Δv_x and $\Delta v'_x$ along direction S are also large. Due to the comparatively larger magnitude of position errors (tens of kilometers), it is not obvious in previous simulation results; the comparatively smaller magnitude of velocity errors (tenths of a meter per second), slight changes will cause comparatively larger relative errors. That explains why relative velocity errors are comparatively larger along direction S in Fig. 2.18.

We can modify results from Monte-Carlo method by the following formula:

$$\begin{cases} \Delta\theta_1 = \arctan[y/(x + r_p)] \\ x' = \sqrt{y^2 + (x + r_p)^2} - r_p, \\ y' = r_p \Delta\theta_1 \end{cases} \begin{cases} \Delta\theta_2 = \arctan[\Delta v_x / (\Delta v_y + v_p)] \\ \Delta v'_x = v_p \Delta\theta_2 \\ \Delta v'_y = \sqrt{\Delta v_x^2 + (\Delta v_y + v_p)^2} - v_p \end{cases} \tag{2.92}$$

where $\Delta v'_x$ and $\Delta v'_y$ are velocity errors after modifications; $\Delta\theta_2$ is an angle between the prediction and real velocities.

Figures 2.21 and 2.22 are result curves of standard deviations after we modified error data using orbit curvature. Relative error of standard deviations of velocity

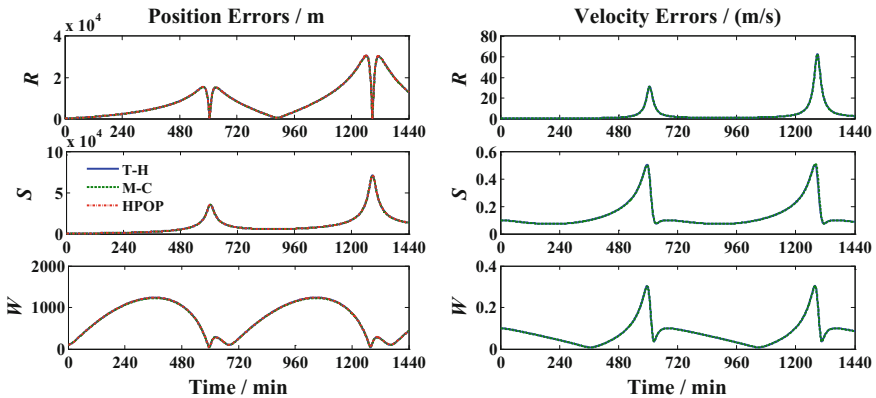


Fig. 2.21 Propagation curve of error standard deviations of object ④ (in RSW coordination system with curvature modifications)

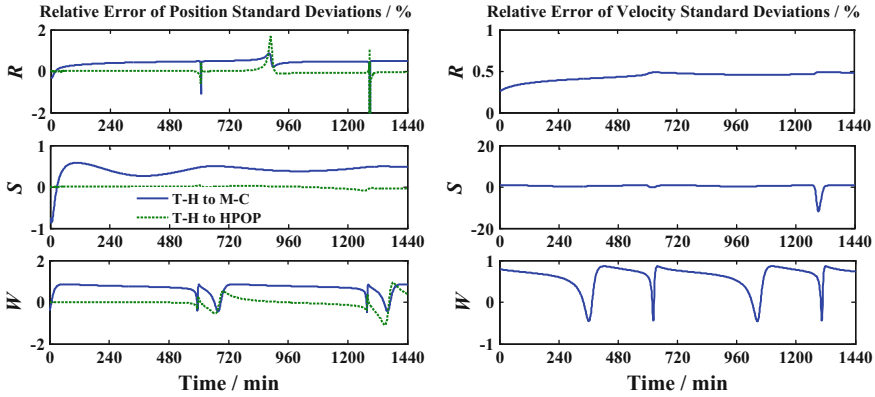


Fig. 2.22 Relative error curve of standard deviations of object ④ (in RSW coordination system with curvature modifications)

errors along direction S reduces to 12%. Velocity error magnitude along direction S is comparatively smaller, so the absolute error corresponding to the relative error of 12% is very small. The two close curves of standard deviations of velocity errors along direction S in Fig. 2.21 also explain that.

We can learn from the derivational process of T-H equation that there are two approximate hypotheses if we describe relative motion. The first one is perturbation force hypothesis, in which the two objects are considered with the same perturbation force and their difference is zero. The second one is linear hypothesis, in which the relative distance of the two objects is considered as a smaller value comparing to their geocentric distances and the equation can be linearized. In the same way, analysis on orbit prediction error propagation with T-H equation is also affected by these two approximate hypotheses. In Figs. 2.27 and 2.21, deviations between T-H and M-C curves reveal deviations caused by linear hypothesis, while deviations between T-H and HPOP curves reveal deviations caused by perturbation force hypothesis. From Fig. 2.22 we can learn that except for the aforementioned particular points, the relative errors between results using T-H method and those of M-C and HPOP methods are usually no more than 2%, which indicates that the precision of analyzing initial errors with T-H equation is adequate.

Due to the relatively large orbital eccentricity and large flight angle of object ④, both RSW and NTW coordinate systems reflect a relatively obvious difference. The propagation and relative error curves of error standard deviations of NTW coordinate system are shown in Figs. 2.23 and 2.24.

From Fig. 2.23 we can learn that error propagation characteristics are different in NTW and RSW coordinate systems. In NTW coordinate system, position errors along direction T and velocity errors along direction N have divergence as time passes, while position errors along direction N have periodic oscillations and velocity errors along direction T have divergence as time passes. In addition, the position error curve along direction R in Fig. 2.17 has sudden changes near the

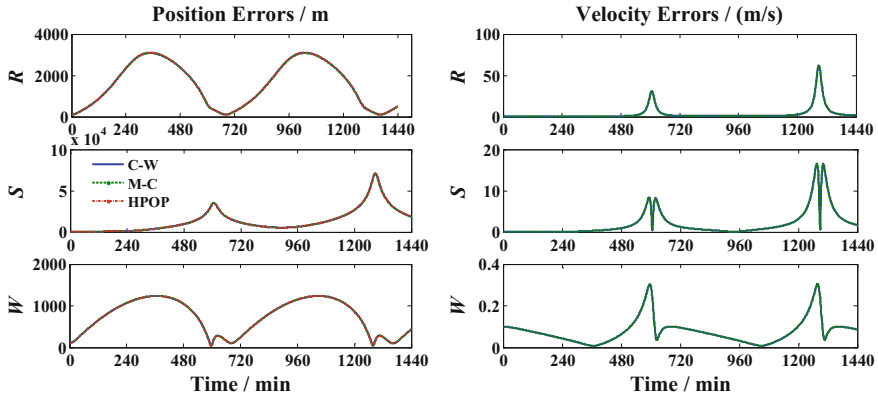


Fig. 2.23 Propagation curve of error standard deviations of object ④ (in NTW coordination system)

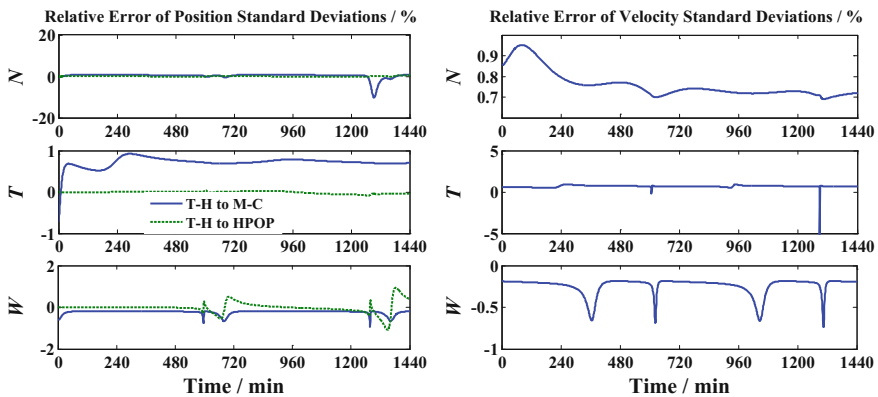


Fig. 2.24 Propagation curve of error standard deviations of object ④ (in NTW coordination system)

perigee and gradually larger amplitude; the position error curve along direction N in Fig. 2.23 is relatively smooth and doesn't increase over time, while velocity errors along direction T have sudden changes near the perigee. We pay more attention to velocity error propagation in practical applications. Therefore, for highly eccentric orbit, we prefer describing errors using NTW coordination system rather than RSW coordination system.

We can make qualitative explanations in Fig. 2.25 for the two characteristics of elliptical orbit errors in RSW coordinate system along direction R (gradually larger amplitude and sudden changes near the perigee). According to theories and practices of orbital dynamics and error propagation [39–41], the long axis of the error ellipsoid is basically identical with the velocity direction, and three principal axis directions are consistent with the three coordinate axis directions in NTW coordinate

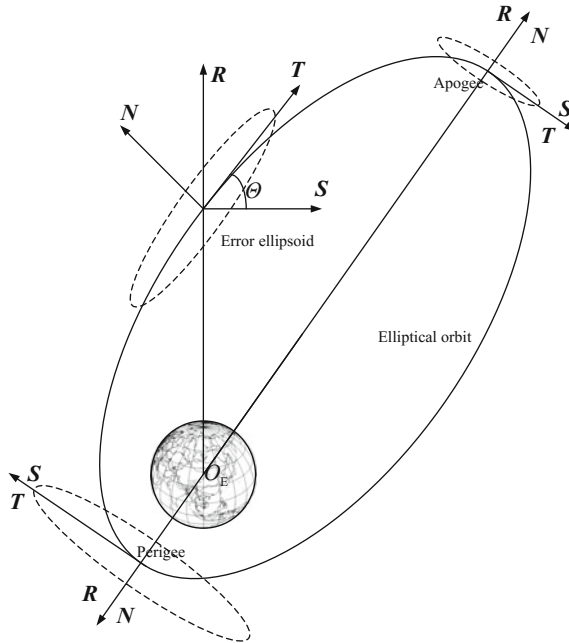


Fig. 2.25 Error ellipsoid of elliptical orbit

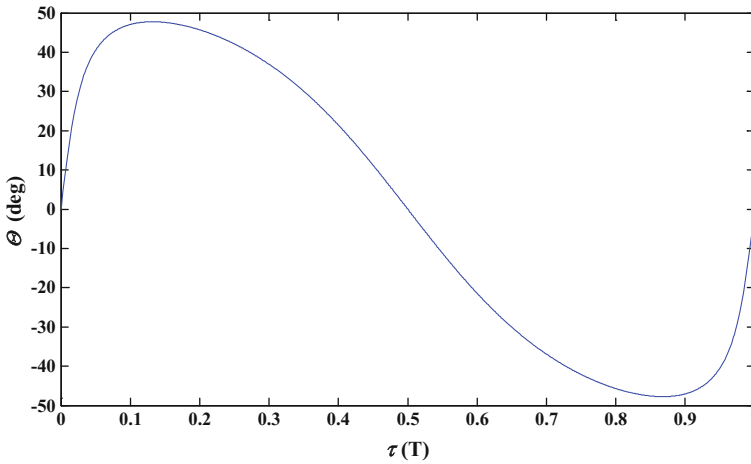


Fig. 2.26 The changing curve of velocity angle θ of object ④ with its passing the Perigee time τ

system as in Fig. 2.26. The angle between Axis S in RSW coordinate system and Axis T in NTW coordinate system is the flight-path angle θ , which is never zero in orbits except for the perigee and apogee. Therefore, the long axis of error ellipsoid

always has projection on Axis R except for the perigee and apogee. As the long axis getting longer over time, the projection enlarges in periodic oscillations, which explains the tendency of long-term position error divergences along direction R .

Figure 2.26 is a changing curve of the flight-path angle of object ④ with its passing the perigee time τ , taking the orbital period as the unit of abscissa. From Fig. 2.26 we can notice that the flight-path angle near the perigee changes from a maximum negative value to zero, then to a maximum positive value. The flight-path angle decreases rapidly to zero near the perigee and the projection of the long axis of error ellipsoid along direction R also decreases rapidly, which explain why there are sudden changes to the errors near the perigee.

2.4.2 Correlation Characterization Based on Geometrical Model

The propagation characteristics of initial error of circular and eccentric orbit are analyzed based on C-W equations and T-H equations respectively in Sect. 2.4.1, which belong to the algebraic model. The negative correlation characteristics of position and velocity error in the case of near-circular orbit are studied based on the geometrical model of relative motion in this section. The application of negative correlation characteristics in the determination of initial error covariance is discussed. The correlation characteristics are validated by historical orbital data.

Orbit prediction error analysis on TLE orbit in Refs. [38, 42] points out that the curve shapes of position errors along S direction and velocity errors along R direction, position errors along R direction and velocity errors along S direction on circular orbit are roughly the same. They are merely inversed around the horizontal axis, which makes them similar. References [38, 42] analyzed this correlation characterization and put forward an assumption that orbit prediction errors are mainly caused by the advancing and falling behind of prediction position on actual orbits, which can explain the generation of long-term error divergences. This assumption only considers the advancing and falling behind of prediction position on actual orbits and does not take the differences between prediction and actual orbits into account. Therefore, it can explain the long-term error divergence characterization, negative correlation characterization of position errors along S direction and velocity errors along R direction, but fails to explain the relation between position errors along R direction and velocity errors along S direction. In this section, we analyze error correlation characterization using geometrical models for relative motion on basis of the references.

Liu has analyzed the correlation characterization of satellite orbital errors [43], derived the relations between satellite position, velocity errors and orbital element errors in RTN frame, provided the quantitative relations between position and

velocity errors in RTN frame on nearly circular orbits and also analyzed the correlation of orbital element errors restrained by position errors in RTN frame.

1. Geometrical Models for Relative Motion

Taking orbital elements of the primary and secondary objects as parameters, the relative dynamics equation is built according to space geometry relation to describe relative motion of space objects, which is called geometrical models for relative motion. Since the geometrical method describes relative motion based on orbital elements, we can apply orbital perturbation theory for the study of the influence of perturbation on relative motion. In the geometrical method, we build relative motion models on primary object orbital coordinate system as well. Schaub [44, 45] came up with a method in which geometrical relationship of relative motion is shown by orbital element differences. We make the relative state vector between two objects in the primary RSW coordinate system as

$$\mathbf{X} = (x, y, z, \dot{x}, \dot{y}, \dot{z})^T \quad (2.93)$$

The orbital element vector of the primary object and the orbital elements difference vector of the secondary to the primary object are

$$\mathbf{e} = (a, e, i, \Omega, \omega, M)^T, \delta\mathbf{e} = (\delta a, \delta e, \delta i, \delta\Omega, \delta\omega, \delta M)^T \quad (2.94)$$

Then the relative position and velocity coordinates in terms of the orbital elements differences can be written as

$$\begin{cases} x = \frac{r}{a}\delta a - a \cos f \delta e + \frac{ae \sin f}{\sqrt{1-e^2}} \delta M \\ y = \frac{2+e \cos f}{1+e \cos f} a \sin f \delta e + r(\delta\omega + \cos i \delta\Omega) + \frac{a^2}{r} \sqrt{1-e^2} \delta M \\ z = r \sin u \delta i - r \cos u \sin i \delta\Omega \\ \dot{x} = -\frac{ne \sin f}{2\sqrt{1-e^2}} \delta a + \frac{a^3}{r^2} \sqrt{1-e^2} n \sin f \delta e + \frac{a^3}{r^2} ne \cos f \delta M \\ \dot{y} = -\frac{3a}{2r} n \sqrt{1-e^2} \delta a + \frac{e+2 \cos f + e(2+e \cos f) \cos^2 f}{(1-e^2)^{3/2}} an \delta e \\ \quad + \frac{ane \sin f}{\sqrt{1-e^2}} (\delta\omega + \cos i \delta\Omega) - \frac{a^3}{r^2} ne \sin f \delta M \\ \dot{z} = \frac{an}{\sqrt{1-e^2}} (\cos u + e \cos \omega) \delta i + \frac{an \sin i}{\sqrt{1-e^2}} (\sin u + e \sin \omega) \delta\Omega \end{cases} \quad (2.95)$$

where

$$r = \frac{a(1-e^2)}{1+e \cos f}, n = \sqrt{\frac{\mu}{a^3}} \quad (2.96)$$

Equation (2.95) depicts the relative position and velocity in RSW coordinate system using the Keplerian orbital element differences, in which $(\dot{x}, \dot{y}, \dot{z})^T$ is the relative velocity error or the temporal derivative of the relative position error. The projection of absolute velocity error into the orbital coordinate system is what more

concerned in error analysis. From the relation of absolute and relative velocities we can get

$$\Delta v_x = \dot{x} - y\dot{f}, \Delta v_y = \dot{y} + x\dot{f}, \dot{f} = \frac{h}{r^2} = \frac{a^2}{r^2} n \sqrt{1-e^2} \quad (2.97)$$

Therefore

$$\begin{cases} v_x = -\frac{ne \sin f}{2\sqrt{1-e^2}} \delta a - \frac{a^2 n \sin f}{r\sqrt{1-e^2}} \delta e - \frac{a^2 n \sqrt{1-e^2}}{r} (\delta \omega + \cos i \delta \Omega) - \frac{a^2 n}{r^2} \delta M \\ v_y = -\frac{an\sqrt{1-e^2}}{2r} \delta a + \frac{e + \cos f + 2e(2+e \cos f) \cos^2 f}{(1-e^2)^{3/2}} an \delta e \\ \quad + \frac{ane \sin f}{\sqrt{1-e^2}} (\delta \omega + \cos i \delta \Omega) \\ v_z = \frac{an}{\sqrt{1-e^2}} (\cos u + e \cos \omega) \delta i + \frac{an \sin i}{\sqrt{1-e^2}} (\sin u + e \sin \omega) \delta \Omega \end{cases} \quad (2.98)$$

For the near-circular orbit ($e \approx 0, r \approx a$), if eliminating terms containing eccentricity e in Eqs. (2.95) and (2.98), we can get

$$\begin{cases} x \approx \delta a - a \cos f \delta e \\ y \approx 2a \sin f \delta e + a(\delta \omega + \cos i \delta \Omega + \delta M) \\ z \approx a(\sin u \delta i - \cos u \sin i \delta \Omega) = a\sqrt{\delta i^2 + \sin^2 i \delta \Omega^2} \sin(u - \phi_z) \\ v_x \approx -an \sin f \delta e - an(\delta \omega + \cos i \delta \Omega + \delta M) \\ v_y \approx -\frac{n}{2} \delta a + an \cos f \delta e \\ v_z \approx an(\cos u \delta i + \sin u \sin i \delta \Omega) = an\sqrt{\delta i^2 + \sin^2 i \delta \Omega^2} \sin(u + \phi_{vz}) \end{cases} \quad (2.99)$$

where the phase angles of position and velocity errors along W direction are

$$\phi_z = \arctan \frac{\sin i \delta \Omega}{\delta i}, \quad \phi_{vz} = \arctan \frac{\delta i}{\sin i \delta \Omega} \quad (2.100)$$

Generally, orbital element errors δe , δi , $\delta \Omega$ and $\delta \omega$ are all small values. They don't change over time (two-body) or change over time slowly (perturbation). The error of semi-major axis δa is largely affected by atmospheric drag which makes its changes faster than other orbital elements aforementioned. If we suppose that the semi-major axis error δa has linear variations over time, we can write it as follows

$$\delta a = \delta a_0 + \delta \dot{a} \cdot \tau, \quad \tau = t - t_0 \quad (2.101)$$

Mean anomaly error δM changes fast over time and have the following relation with initial semi-major axis error δa_0 and semi-major axis rate $\delta \dot{a}$:

$$\delta M = \delta M_0 - \frac{3n}{2\alpha} \left(\delta a_0 \tau + \frac{\delta \dot{a}}{2} \tau^2 \right) \quad (2.102)$$

Substitute Eqs. (2.101) and (2.102) into the terms in the orbital plane of Eq. (2.99), we can get

$$\begin{cases} x \approx \delta\dot{a}\tau - a \cos f \delta e + \delta a_0 \\ y \approx \left(-\frac{3}{4}n\delta\dot{a}\tau^2 - \frac{3}{2}n\delta a_0\tau \right) + 2a \sin f \delta e + a(\delta\omega + \cos i \delta\Omega + \delta M_0) \\ v_x \approx \left(\frac{3}{4}n^2 \delta\dot{a}\tau^2 + \frac{3}{2}n^2\delta a_0\tau \right) - an \sin f \delta e - an(\delta\omega + \cos i \delta\Omega + \delta M_0) \\ v_y \approx -\frac{n}{2}\delta\dot{a}\tau + an \cos f \delta e - \frac{n}{2}\delta a_0 \end{cases} \quad (2.103)$$

Equation (2.103) reveals that there are linear term τ and quadratic term τ^2 of time in position errors along S direction and velocity errors along R direction, which will have fast divergences as time passes. In position errors along R direction and velocity errors S along direction, there are only linear term τ , initial semi-major axis error δa_0 and eccentricity error δe . Because semi-major axis change rate $\delta\dot{a}$ has a small magnitude, the linear term of time which takes it as parameter has smaller changes over time and not obvious long-term divergences. The position errors and velocity errors along W direction are defined by direction errors of the orbital plane (the inclination error δi and the right ascension of ascending node error $\delta\Omega$), and oscillates with the orbital period. The periodicity of position and velocity errors on the orbital plane (along R and S direction) is caused by eccentricity error δe and its period is also orbital period.

2. Correlation of Position and Velocity Errors

(1) Relations of position and velocity errors along R and S direction

Equation (2.103) indicates that the expression of position error y along S direction and velocity error v_x along R direction is composed by three parts for near-circular orbits:

- (1) secular term (up to second order) caused by initial semi-major axis error δa_0 and its change rate $\delta\dot{a}$;
- (2) periodic term caused by eccentricity error δe ;
- (3) constant term caused by argument of perigee error $\delta\omega$, right ascension of the ascending node error $\delta\Omega$ and initial mean anomaly error δM_0 .

Position error y along S direction and velocity error v_x along R direction have the following approximate relation:

$$\frac{v_x}{y} \approx -n = -\frac{2\pi}{T} \quad (2.104)$$

Equation (2.103) indicates that the expression of position error x along R direction and velocity error v_y along S direction is also composed by three parts for near-circular orbits:

- (1) one-order secular term caused by semi-major axis change rate $\delta\dot{a}$;
- (2) periodic term caused by eccentricity error δe ;
- (3) constant term caused by initial semi-major axis error δa_0 .

Position error x along R direction and velocity error v_y along S direction have the following approximate relation:

$$\frac{v_y}{x} \approx -n = -\frac{2\pi}{T} \quad (2.105)$$

Equations (2.104) and (2.105) represent the approximate relations of position and velocity errors along R and S direction for near-circular orbit. For near-circular orbit, quantitative relation is obvious between position and velocity errors along R and S direction:

- (1) For magnitude, the R and S components of velocity errors are $n = 2\pi/T$ times than the S and R components of position errors. For LEO, its period $T \approx 100\text{min}$, n is about 1/1000; for MEO, its period $T \approx 12\text{h}$, n is approximately 1/7000; for GEO, its period $T \approx 24\text{h}$, n is approximately 1/14,000.
- (2) For sign, the sign of the R and S components of velocity errors are opposite to the S and R components of position errors, which explain that the curve shapes of position errors along S direction and velocity errors along R direction on circular orbit are approximately the same. They are merely inversed around the horizontal axis, which makes them similar.

(2) Relations of position and velocity errors along W direction

Position and velocity errors along W direction have periodic oscillations, and their oscillation functions are

$$z \approx a\sqrt{\delta i^2 + \sin^2 i \delta \Omega^2} \sin(u - \phi_z), v_z \approx an\sqrt{\delta i^2 + \sin^2 i \delta \Omega^2} \sin(u + \phi_{v_z}) \quad (2.106)$$

Equation (2.100) infers that phase angles of position and velocity errors meet:

$$\phi_z + \phi_{v_z} = \frac{\pi}{2} \quad (2.107)$$

Therefore, the v_z expression in Eq. (2.106) can be written as

$$v_z \approx na\sqrt{\delta i^2 + \sin^2 i \delta \Omega^2} \sin\left(u + \frac{\pi}{2} - \phi_z\right) \quad (2.108)$$

The relation of position error z and velocity error v_z along W direction is

$$v_z(u) = nz\left(u + \frac{\pi}{2}\right) = \frac{2\pi}{T}z\left(u + \frac{\pi}{2}\right) \quad (2.109)$$

where $v_z(\cdot)$ and $z(\cdot)$ denote that position and velocity errors are functions of the argument of latitude u . Equation (2.109) represents the approximate relation of position and velocity errors along W direction on near-circular orbit. For amplitude, velocity errors are $n = 2\pi/T$ times than position error. For phase, velocity errors are $\pi/2$ ahead of position errors.

3. Negative Correlation Characteristic of Error Covariance Matrix

From the last section we can know that position errors along S direction and velocity errors along R direction are in approximate quantitative relation as in Eq. (2.104), and position errors along R direction and velocity errors along S direction are also in approximate quantitative relation as in Eq. (2.105). This characteristic is actually the negative correlation of position and velocity errors, which reflects in error covariance matrix, i.e. the negative correlation characteristic of covariance matrix. We can use a correlation coefficient to express this inverse relationship and approximate characterization. The definition of correlation coefficient ρ is [46]

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\sigma^2(X)\sigma^2(Y)}} \quad (2.110)$$

where $\text{cov}(\cdot, \cdot)$ and $\sigma^2(\cdot)$ are covariance and variance respectively, correlation coefficient is dimensionless.

The absolute value of correlation coefficient ρ is not larger than 1, i.e. $|\rho| \leq 1$. $|\rho|$ getting larger indicates that X and Y have closer connections for linear relation; especially when $|\rho| = 1$, X and Y are in linear relation with probability 1. When $\rho > 0$, X and Y have same tendencies; when $\rho < 0$, X and Y have opposite tendencies. When ρ is near 1, positive linear relation is between two random variables; when ρ is near -1 , negative linear relation is between two random variables (negative correlation); when ρ is near 0, no linear relations are between two random variables.

The expression of orbital error covariance matrix $P(t)$ using C-W equations is got in Sect. 2.4.1. The correlation coefficients could be gotten through the following equation:

$$\rho_{x,v_y} = \frac{\text{cov}(x, v_y)}{\sigma(x)\sigma(v_y)} = \frac{P_{1,5}}{\sqrt{P_{1,1}P_{5,5}}}, \quad \rho_{y,v_x} = \frac{\text{cov}(y, v_x)}{\sigma(y)\sigma(v_x)} = \frac{P_{2,4}}{\sqrt{P_{2,2}P_{4,4}}} \quad (2.111)$$

The initial covariance matrix is then propagated, the correlation coefficients are obtained by Eq. (2.111). Figures 2.27, 2.28 and 2.29 illustrate the curves of two correlation coefficients ρ_{x,v_y} and ρ_{y,v_x} of three objects over time. Note that these two correlation coefficients in initial covariance matrix are both zero.

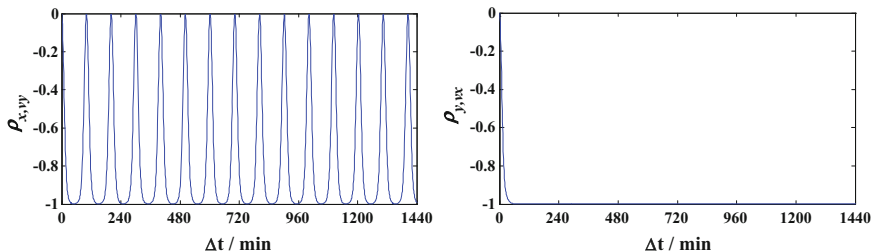


Fig. 2.27 Correlation coefficients of LEO object based on C-W equations

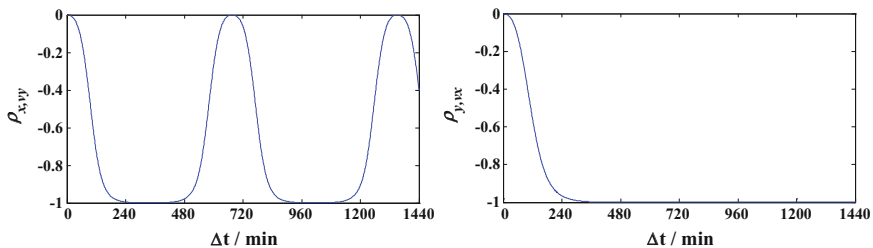


Fig. 2.28 Correlation coefficients of MEO object based on C-W equations

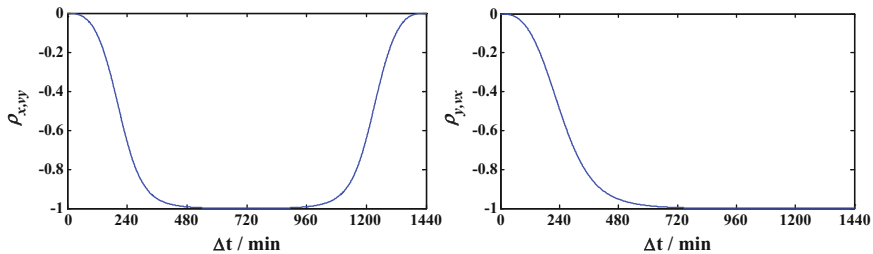


Fig. 2.29 Correlation coefficients of GEO object based on C-W equations

Figures 2.27, 2.28 and 2.29 show that the correlation coefficient ρ_{y,v_x} between position error along S direction and velocity error along R direction changes from 0 to -1 very quickly after the beginning of the prediction (within half an orbital period), which indicates that there is strong negative correlation between position error along S direction and velocity error along R direction. The correlation coefficient ρ_{x,v_y} of position error along S direction and velocity error along R direction is permanently not larger than zero and has oscillations from 0 to -1 , with the orbital period. There is also strong negative correlation between position error along R direction and velocity error along S direction.

Table 2.6 Epoch and initial orbital elements

Epoch (UTCG)	a/km	e	$i/(\circ)$	$\Omega /(\circ)$	$\omega/(\circ)$	$M/(\circ)$
2012-07-01 00:00:00.00	7068.920954	0.002805	98.141	256.737	96.061	180.974

Table 2.7 Initial position and velocity covariance matrix

		x	y	z	v_x	v_y	v_z
$\rho = 0$	x	25					
	y		100				
	z			25			
	v_x				4×10^{-4}		
	v_y					1×10^{-4}	
	v_z						1×10^{-4}
$\rho = -0.9$	x	25				-0.045	
	y		100		-0.18		
	z			25			
	v_x		-0.18		4×10^{-4}		
	v_y	-0.045				1×10^{-4}	
	v_z						1×10^{-4}
$\rho = 0.9$	x	25				0.045	
	y		100		0.18		
	z			25			
	v_x		0.18		4×10^{-4}		
	v_y	0.045				1×10^{-4}	
	v_z						1×10^{-4}

4. Application of Correlation Characteristic in Error Propagation

One of the applications of the negative correlation characteristic of orbital error is the determination of initial error covariance matrix. A LEO object is taken as an example to explain. Table 2.6 shows the epoch and initial orbital elements.

Three conditions are considered when determine the initial covariance matrix: (1) the position and velocity errors are non-correlated; (2) the position and velocity errors are in negative correlation, the correlation coefficients $\rho_{x,v_y} = \rho_{y,v_x} = -0.9$; (3) the position and velocity errors are in positive correlation, the correlation coefficients $\rho_{x,v_y} = \rho_{y,v_x} = 0.9$. Table 2.7 shows the initial position and velocity covariance matrix under these three conditions.

We use high precision orbit propagation (HPOP) models to perform orbit propagation and covariance propagation. Under the three initial covariance conditions, we use the STK/HPOP covariance propagation module to perform initial error covariance matrix prediction and get changing curves of error standard deviations of one-day and ten-day predictions in RSW coordination system as in

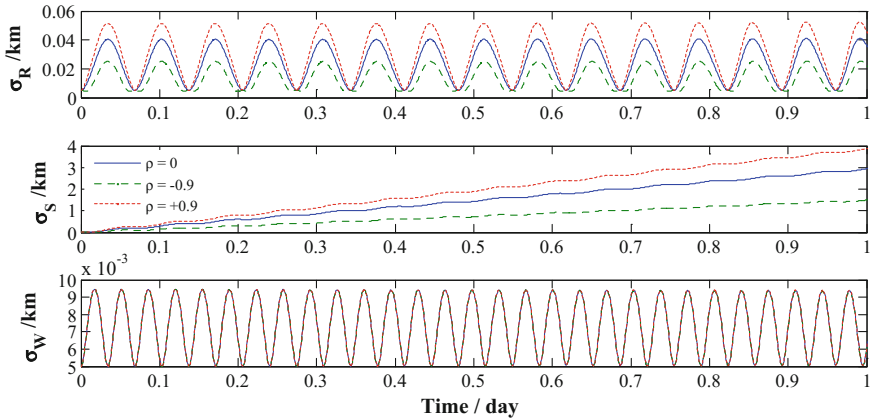


Fig. 2.30 Propagation of initial covariance (1 day)

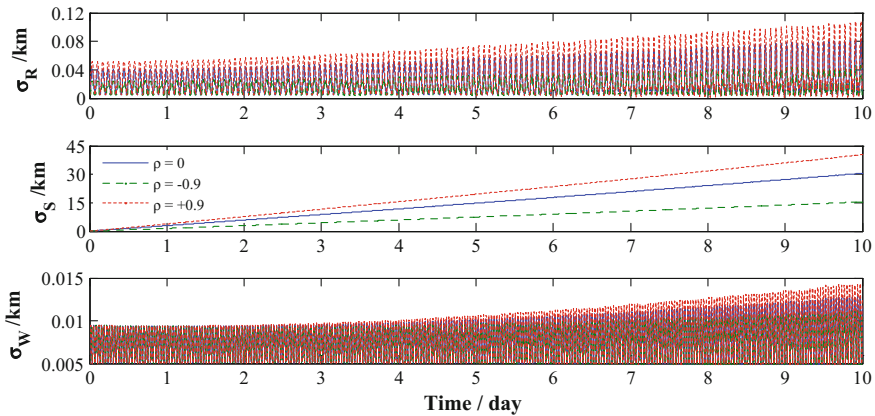


Fig. 2.31 Propagation of initial covariance (10 day)

Figs. 2.30 and 2.31. These figures indicate that the consideration of negative correlation characteristic of position and velocity errors when determining initial covariance can obviously reduce the divergence of initial error.

The negative correlation characteristic of position and velocity errors is derived under the assumption that the orbit is near-circular and the orbital element error are small values. If the errors of near-circular orbit do not meet negative correlation characteristics, the orbital element error is not a small value and orbit prediction errors must have serious divergence. Due to the internal relations among position and velocity errors at all directions, we must consider the relations when defining initial covariance matrix to make sure that the parameters are self-consistent and reasonable.

References

1. Liu L (2000) Orbit theory of spacecraft. National Defense Industry Press, Beijing
2. Vallado DA (2004) Fundamentals of astrodynamics and applications, 2nd edn. Microcosm Press, El Segundo
3. CCSDS Secretariat. Draft recommendation for space data system standards: Conjunction Data Message. Draft recommended standard CCSDS 508.0-R-1, 2012
4. Li JS (2003) Orbit determination of spacecrafts. National Defense Industry Press, Beijing
5. Gelb A, Warren RS (1972) Direct statistical analysis of nonlinear system—CADET. In: AIAA guidance and control conference, pp 72–875
6. Taylor JH, Price CF (1974) Direct statistical analysis of missile guidance systems via CADET. Analytic Sciences Corporation, Massachusetts
7. Taylor JH (1975) Handbook for the direct statistical analysis of missile guidance system via CADET. Analytic Sciences Corporation, Massachusetts
8. Gelb A, Kasper J, Price C et al (1974) Applied optimal estimation. MIT Press, Cambridge, MA
9. Liang LB, Luo YZ, Zhang J et al (2010) Rendezvous—phasing errors propagation using quasi—linearization method. In: AIAA2010-7594. AIAA guidance, navigation, and control conference, Toronto, Ontario
10. Liang LB, Luo YZ, Xing JJ et al (2010) Precision analysis of nonlinear rendezvous by covariance analysis description equation technique. Syst Eng Electron 32(9):1977–1981
11. Vallado DA (2005) An analysis of state vector propagation using differing flight dynamics programs. In: AAS 05-199. AAS/AIAA space flight mechanics conference, Copper Mountain, Colorado
12. Vallado DA (2007) A preliminary analysis of state vector prediction accuracy. In: AAS 07-358. AAS/AIAA space flight mechanics meeting, Sedona, Arizona, 28 January–01 February 2007
13. Chan JC, Navarro D (2001) Comparisons of NORAD two-line elements with INTELSAT orbital elements. In: Proceeding of the third european conference on space debris, Darmstadt, Germany
14. Kelso TS (2007) Validation of SGP4 and IS-GPS-200D against GPS precision ephemerides. In: AAS 07-127. AAS/AIAA space flight mechanics meeting, Sedona, Arizona, 28 January–01 February 2007
15. Boyce WH (2004) Examination of NORAD TLE accuracy using the Iridium constellation. Spaceflight mechanics 2004. Advances Astron Sci 119:2133–2142
16. Snow D, Kaya D (1999) Element set prediction accuracy assessment. Astrodynamics 1999. Advances Astron Sci 103:1937–1958
17. Muldoon AR, Elkaim GH (2008) Improved orbit estimation using GPS measurements for conjunction analysis. In: 21st international technical meeting of the satellite division of the institute of navigation, IONGNSS 2008
18. Zhang JH (1990) Martingale representation of nonlinear system error covariance matrix: an annotation for CADET. J Spacecraft TT&C Technol 1:1–5
19. Han L (2008) Orbit determination and error analysis with space and ground cooperation for LEO space surveillance. National University of Defense Technology, Changsha
20. Junkins J, Akella M, Alfried K (1996) Non-Gaussian error propagation in orbital mechanics. J Astron Sci 44(4):541–563
21. Yang LP, Zhu YW, Huang H (2010) Planning and control for relative motion trajectory of spacecraft. National Defense Industry Press, Beijing
22. Zhang YK (2002) Study on the technique of satellite formation flying dynamics and control. National University of Defense Technology, Changshang
23. Meng YH (2006) Research on control and application of LEO spacecraft formation flying. National University of Defense Technology, Changsha

24. An XY (2006) Dynamics and application of spacecraft formation flying in eccentric orbits. National University of Defense Technology, Changsha
25. Li JF, Gao YF, Baoyin HX (2002) Study on satellite formation flying dynamics and control. *Mech Eng* 24(2):1–6
26. Gao YF, Baoyin HX, Li JF (2002) Dynamics behavior and simulation of relative trajectories of satellite formation flying. *J Tsinghua Univ (Sci Technol)* 42(4):458–461
27. Hill GW (1878) Researches in Lunar theory. *Am J Math* 1:5–26
28. Clohessy WH, Wiltshire RS (1960) Terminal guidance system for satellite rendezvous. *J Aerosp Sci* 27(9):653–674
29. Yamanaka K, Ankersen F (2002) New state transition matrix for relative motion on an arbitrary elliptical orbit. *J Guid Control Dyn* 25(1):60–66
30. Yue XK, Yuan YX (2011) Transfer matrix for relative dynamics in elliptic orbit. *Chinese Space Sci Technol* 1:42–47
31. Xing JJ (2007) Study on formation design and stationkeeping of spacecraft formation flying. National University of Defense Technology, Changsha
32. Xing JJ, Tang GJ, Xi XN et al (2007) Satellite formation design and optimal station keeping considering nonlinearity and eccentricity. *J Guid Control Dyn* 30(5):1523–1528
33. Xing JJ, Li HY, Tang GJ et al (2006) Periodic relative motion condition for satellites formations considering nonlinearity. *J Astronaut* 27(3):359–362
34. Xing JJ, Tang GJ, Xi XN et al (2006) Nonlinear and periodic relative motion in spacecraft formations in eccentric orbits. *J Tsinghua Univ (Sci Technol)* 46(8):1462–1465
35. Xing JJ, Tang GJ, Li HY (2006) New method of the analytic periodic solution for spacecraft formation in elliptical orbits. In: 57th international astronomical congress, Spain
36. Xing JJ, Li HY, Tang GJ et al (2006) Periodic-relative rotating orbits design for spacecraft formations by numerical optimization. *J National Univ Defense Technol* 28(1):13–16
37. Xing JJ, Hao XN, Wang W et al (2003) Relative states autonomous determination of two satellites formation. *J Astronaut* 24(3):254–258
38. Chen L, Han L, Bai XZ et al (2010) Orbital dynamics and error analysis of space target. National Defense Industry Press, Beijing
39. Chan FK (2008) Spacecraft collision probability. The Aerospace Press, El Segundo
40. Alfriend KT, Akella MR, Frisbee J et al (1999) Probability of collision error analysis. *Space Debris* 1:21–35
41. Ghrist RW, Plakalovic D (2012) Impact of non-Gaussian error volumes on conjunction assessment risk analysis. In: AIAA 2012-4965. AIAA/AAS astrodynamics specialist conference, Minneapolis, Minnesota
42. Bai XZ (2008) Research on collision probability in space objects collision detection. National University of Defense Technology, Changsha
43. Liu ZG (2011) Relativity of satellite orbit error. *J Spacecraft TT&C Technol* 30(5):45–49
44. Schaub H, Alfriend KT (2002) Hybrid Cartesian and orbit element feedback law for formation flying spacecraft. *J Guid Control Dyn* 25(2):387–393
45. Schaub H (2004) Relative orbit geometry through classical orbit element differences. *J Guid Control Dyn* 27(5):839–848
46. Ma ZH (2000) Handbook of modern applied mathematics: volume for probability statistics and stochastic process. Tsinghua University Press, Beijing

Chapter 3

Orbital Error Analysis Based on Historical Data

For operational spacecraft, the owner/operator is in charge of the orbit determination. They process the measurements from the tracking stations and come up with a best estimate of the object's orbit, measurement residuals, and the initial covariance matrix. The initial covariance matrix can be propagated by means of the transition matrix. The transition matrix contains the derivatives of the state vector at a given position with respect to the dynamic parameters (initial state and force model parameters). It depends on the model used for the propagation of the orbit.

For most of the non-cooperative objects, due to large amount and insufficiency of high-quality observations, they must be maintained by mean elements and analytical orbital model in most space object catalogs. The covariance matrix is not released in most analytical space object catalogs. For those object only historical orbital element data available, the error statistical fitting is a feasible method, which directly obtains evolution characteristics of prediction error with respect to time and other parameters.

This chapter characterizes the periodicity of orbital prediction error and introduces Poisson series for error fitting. Note that TLE's prediction accuracy is relatively high around its epoch, the prediction states in an orbital period centered at its epoch are taken as reference states. The comparison is conducted in a whole period so that the residuals can reflect the periodicity of TLE error. Poisson series rather than polynomial series is taken as error fitting function, which can describe variation of error with respect to propagation duration and on-orbit position of objects. The error analysis and fitting method can not only handle TLE sets, but also handle any other sequential analytical or numerical orbital elements. Currently, the most complete source of historical orbital data available to the public on the full catalog of space objects are NORAD's TLEs. Therefore, TLEs are taken as example to address the method.

3.1 Overview of Error Analysis and Application Based on Historical Data

For most of the non-cooperative objects (such as inactive spacecraft, rocket bodies, and space debris), due to their large amount and insufficiency of high-quality observations, they must be maintained by mean elements and analytical orbital model in most space object catalogs. For example, orbital data used most widely in conjunction analysis are North American Aerospace Defense Command (NORAD)'s Two-Line Elements Sets (TLE). TLEs are mean elements and the associated orbital propagator SGP4/SDP4 is analytical. The covariance matrix is not released in most analytical space object catalogs. No a priori accuracy is given with element sets, and the propagation models to be used with are also an additional source of inaccuracy. The risk assessment faces the problem that no covariance information is available for the orbital element data.

U.S. Strategic Command (USSTRATCOM) has warned via Space-track website that TLEs should not be used for conjunction assessment predictions. Even though, it is advantageous to know the prediction accuracy and error evolution of TLEs for their widely used application in pre-filter, closest approach analysis, and orbital anomaly detection. Therefore, a practical TLE accuracy assessment and error covariance analysis technique is required.

Many studies have been carried out on the accuracy assessment and error covariance analysis. The orbital prediction accuracy can be examined by comparing TLE predicted ephemerides to high-precision ephemerides for operational satellites with available high-precision ephemerides. This method has been applied to INTERLSAT fleet [1], GPS constellation [2], and Iridium constellation [3]. Snow and Kaya compared the accuracy of the TLE data to the observation data used to generate the TLE [4]. This kind of accuracy assessment is not practical for most users because the observation data and high-precision ephemerides are not generally available. Additionally, each object's orbit error has its own characteristic and should be assessed individually, so a few satellites with high-precision ephemerides cannot totally represent all the objects in similar orbits.

For those object only historical orbital element data available, the error statistical fitting is a feasible method. The basic idea is comparing orbital states (predicted state or estimated state from orbital data processing) at or around TLE's epoch with the predicted states using prior TLEs and SGP4 propagation model. The residuals are statistically analyzed to obtain error information and temporal evolution of statistical quantities such as mean and covariance. The assumption of this method is that the catalogue of object is steady and the series orbital data have approximate precision level.

Most of the methods developed so far can be classified into (1) initial error analysis and (2) error evolution analysis. Besides, the historical orbital data could also promote the long-term prediction precision of TLE.

3.1.1 Initial Covariance Analysis Based on Historical Data

Historical-data-based initial error analysis obtains the initial covariance matrix at the epoch of each TLE. Krag and Flohrer et al. coped with this issue by introducing pre-defined look-up tables for the initial covariance that are sorted by eccentricity, perigee altitude and inclination [5, 6].

The initial covariance was obtained from comparing states derived directly from the TLE data with states resulting from an orbit determination using pseudo-observations derived from TLE data by a numerical propagation. Initial covariance was applied to the propagation of the state and covariance. The obtained covariance information reflects both the limitations of the TLE (SGP4) orbit model and the limitations of orbit determination and propagation accuracy.

Wang [7] obtained errors analysis of classified orbits and typical orbit correlated with SSO by using SGP4/SDP4 model and TLE data, she also discussed B^* in TLE data and tried to modify it. Results show that after B^* modification, orbit prediction precision becomes better.

Kelso [2] pointed out that error characteristics for objects in similar orbits could be considerably different, so the error characteristics of each object should be determined independently. In addition, one should propagate the state and covariance forward or backward by numerical orbital model to obtain covariance matrix at specific time (TCA for instance). That is complex and computationally inefficient.

3.1.2 Covariance Propagation Function Fitting Based on Historical Data

Differing from initial error analysis, error evolution analysis directly obtains evolution characteristics of prediction error with respect to time and other parameters. Each space object has an associated time series of TLE data, this time series forms the basic data set available. The basic method is comparing states at or around TLE's epoch with the predicted states using prior TLEs and SGP4 propagation model. The residuals are statistically analyzed to obtain error information and temporal evolution of statistical quantities such as mean and covariance.

Peterson [8] attempted to generate general error information by relying solely upon the publicly available TLE data. The program used in constructing the error is COVGEN (COVariance GENeration). The COVGEN program produced an estimate of the error and error growth for a given object. It processed a time series of TLEs for a specific object and found the differences in satellite-based coordinate system, performed a quadratic least squares estimate to determine the error growth. Due to uniqueness of its solution and independence on observation residuals, COVGEN provided an independent measure of consistency of the TLEs covariances produced by the true observation residuals. In addition, the COVGEN

methodology is adaptable enough that any type of ephemeris data can be used and will produce a covariance comparable in quality to the input data.

Deguine [9] presented an approach that allows statistically evaluating the accuracy of orbital position and velocity and B^* parameter by using TLE and the associated propagation models. This accuracy information allows characterizing the temporal evolution of the TLE covariance matrix for fixed orbit propagation durations. The technique relies on the identification of a temporal error model based on the calculation of differences between the propagation of TLE and the reference over a fixed delay.

Oswiler [10] estimated covariance matrix and autocorrelation function of state vector using only publicly available TLEs. He also evaluated the variances and covariances for satellites in different orbits as they change over time. In addition to analyzing the consistency of the covariance estimates, the dependability of propagated TLE state vectors with respect to the length of prediction was evaluated by using the autocorrelation relationship.

Legendre [11–13] indicated that it's rough to approximate the TLE error distribution by a normal law. They approximated the TLE error distribution by a Gaussian mixture and used the standard Expectation Maximization (EM) algorithm to determine describing parameters of the Gaussian laws. They handled one-dimensional [11], three-dimensional [12], and propagation duration depended problem [13] by mixture of Gaussian laws.

Matney [14] and Hirose [15] also studied the accuracy assessment and covariance estimation problem based on historical orbital data.

Current widely used covariance modeling methods conducted pair-wise differencing at epochs of TLEs. The fitting function was temporal polynomial, which cannot characterize periodicity of orbital error. Orbital prediction errors depend not only on propagation duration but also on on-orbit position, especially for highly eccentric orbit (HEO), as they experience more complex perturbations, including atmospheric drag, solar radiation pressure, and third-body gravity [10]. HEO's perturbation situation varies from perigee to apogee remarkably, causing the obvious variation of error from perigee to apogee and the uncertainty ellipsoid to breathe as the object operates on its orbit. Therefore, periodicity of orbital prediction error must be characterized to obtain more accurate and reasonable covariance information for HEO objects. Most TLEs are released at the epoch when the objects are passing its orbit's the ascending node where the argument of latitude is about zero, so the residuals at the epoch cannot reflect the periodic characteristics of TLE error.

Some of the previous studies have paid attention to the periodic characteristic of orbital prediction error. For instance, Legendre [13] have considered the dependence on on-orbit position (or argument of latitude) of TLE error and introduced a time delay δt corresponding to the on-orbit position. Legendre [13] must conduct close approach analysis first to get the TCA of the conjunction event and corresponding argument of latitude (AOL) of dangerous object. They introduced an additional AOL-related time delay δt into epoch of TLE so that all the TLE errors are generated at specific AOL. Due to mathematical complexity, they had to restrict

Gaussian mixture to a 2-D mixture and project the 3-D TLE errors into the conjunction plane, so the accuracy estimation was directly related to specific conjunction geometry. Due to the accuracy estimation is linked to a specific AOL and conjunction geometry, only after close approach analysis could the TLE accuracy estimation be done, this might be disadvantageous to timely conjunction risk assessment.

The historical-data-based methods without high precision ephemerides virtually test the consistency between subsequent TLEs. By studying TLE accuracy and consistency of objects whose high precision orbit ephemerides was available, some authors have pointed out that TLE consistency matches very well the real accuracy. Kelso [2] pointed out that TLE consistency analysis does reasonably approximate the true error of a TLE prediction, both in propagation time, direction, and overall magnitude. In spite of that this affirmation could not be generalized for all the objects contained on the catalogue, the consistency analysis using pair-wise comparison maybe is the only feasible way for approximating orbit accuracy for those object only historical orbital element data available. With the development of space surveillance capability, this situation will change in future.

3.1.3 Promotion of Long-Term Precision Using Historical Data

The historical orbital data of space objects could not only be used to the quantitative analysis of prediction error, but also to the promotion of prediction precision.

The only source of knowledge for the majority of space debris is the publicly available two-line element (TLE) sets. But predictions based on TLEs using the associated analytic propagator (SGP4) are not sufficiently accurate to warrant maneuvering to avoid potential collisions. A single TLE and the high-precision numerical propagator cannot achieve high orbit prediction accuracy all the same. To address this accuracy problem, several methods to improve the propagation accuracy for non-maneuvering orbital objects using only historical TLEs as input data are investigated.

Levit [16] describes a method to increase the predictive accuracy for orbital objects using only historical TLE data, such that it enables operational conjunction assessment for collision avoidance. The method is essentially to use state vectors derived from successive TLEs around their epoch as pseudo-observations, and to fit an orbit to these pseudo-observations using a high-precision numerical propagator and traditional batch least-squares differential correction. The fitted orbit is then propagated into the future using the same high-precision orbit propagator, as seen in Fig. 3.1. The method leads to a predicted range error that increases at a typical rate of 100 m per day, approximately a 10-fold improvement over individual TLE's propagated with their associated analytic propagator (SGP4). The limitation of this

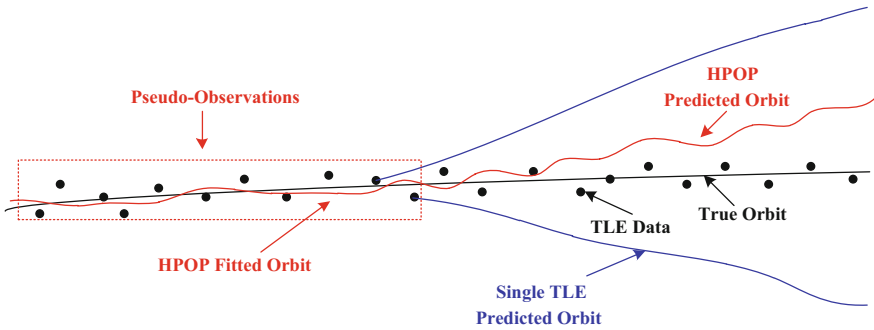


Fig. 3.1 The scheme for increasing long-term precision using historical data

method is that the least-squares smoothing can minimize the effects of variance but cannot remove the bias in a sequence of TLEs.

Muldoon [17] presents an approach to orbital propagation using data-driven modeling techniques to predict future orbital state. The model fitting system for each of the TLE parameters is trained using the historical data available for an object. This model is evaluated at the future point to estimate the future value. The model fitting system uses blending of several window-optimized polynomial models of different orders for estimating epoch time, and a polynomial plus optimal periodic components for all other parameters. This approach achieves accuracy comparable to or better than SGP4.

General, the collision probability is computed by assuming that the covariances characterizing the position uncertainties of two objects are independent. However, both objects suffer kinds of perturbations simultaneously. For instance, the uncertainty in the atmospheric model (due to unpredictable variations in solar activity) is the main error source for objects in low-Earth orbits (LEO). Because of the coherent nature of atmospheric drag, the independent assumption is not necessarily true, especially for the down-range uncertainty. If the solar activity goes up, both objects will have increased drag. Likewise, if the solar activity goes down, both objects will have decreased drag. Therefore, in order to compute the collision probability to maximum accuracy, it will be necessary to include the correlation effect. Matney [14] uses two-line element set data to demonstrate the covariance correlations and estimates how much this may affect the accuracy of the calculated collision probability.

It is usually assumed that the position uncertainties are uncorrelated between the two objects - that is, there is no cross correlation contribution to the relative position covariance. Coppola examines the sensitivity of the computed collision probability to non-zero cross correlation [18]. In certain cases, small but physically realistic cross correlation may lead to large variations in collision probability.

Table 3.1 List of space objects selected for the study

Orbit type	Catalog number	Satellite name	Inclination /°	Perigee altitude/km	Apogee altitude/km	TLE sets
LEO	22675	Cosmos-2251	74.04	776	799	612
MEO	32711	GPS-62	55.55	20,096	20,267	679
HEO	28544	SLOSHSAT	7.08	245	35,108	508
GEO	25010	TelStar-10	0.06	35,771	35,802	821

3.2 Generation of Residual Data Considering Periodicity

3.2.1 Selection of Space Objects

The initial task is to select a number of satellites as examples to demonstrate and validate the error analysis and fitting method. The satellites selected form a spectrum including low Earth orbit (LEO), medium Earth orbit (MEO), highly eccentric orbit (HEO), and geostationary orbit (GEO). The selected satellites and some of their orbital parameters are summarized in Table 3.1. Numbers of historical TLE sets of each object are also included in the table. All TLEs used in this study are obtained from www.space-track.org. Each object in Table 3.1 is referred as its orbital type in the following text for convenience. We will refer Cosmos-2251 as LEO object and SLOSHSAT as HEO object for instance.

3.2.2 Generation of Residual Data

Error analysis and fitting require the residual data between predicted states (position and velocity) and reference states. The straightforward approach to generate residual data is calculating the differences between the predicted state vector and the “real” state vector at specific times.

$$\Delta \mathbf{X} = \mathbf{X}_{\text{pred}} - \mathbf{X}_{\text{real}} \quad (3.1)$$

However, the “real” state vector \mathbf{X}_{real} is always unknown in practice, and must be approximated. The only available information in this study is the TLE data itself, so some kind of predicted value must be used to approximate the ‘real’ value. Generally, the prediction error around TLE’s epoch is relatively small (but not always the minimum), so in the prior works the state vectors derived from each TLE at their epoch are taken as approximation of “real” states. Most TLEs are released when objects are passing its orbit’s ascending node where the argument of latitude is about zero, so the residuals at epoch cannot reflect the periodic characteristics of TLE error. Considering that TLE’s prediction errors in an orbital period centered at its epoch are relatively small, this study takes the predicted states

in this period as reference states. The predicted states of prior TLEs to this interval are used as prediction. Differencing the predictions and references, we can obtain residual data.

For the convenience to discuss, these nomenclatures are used below:

- t_i the epoch of i th TLE set, where $i = 0, 1, \dots, N_{\text{tle}} - 1$, N_{tle} is the number of all TLE sets;
- $t_{i,k}$ the k th time point in the time interval centered at the i th TLE's epoch;
- t_T the threshold of prediction time in the error analysis, e.g. 7 days, 15 days, or 30 days, et al.
- $\Delta t_{i/j,k}$ the temporal difference between the time point $t_{j,k}$ and the epoch of i th TLE set, $\Delta t_{i/j,k} = t_{j,k} - t_i$;
- $X(t_i/t_{j,k})$ the predicted state vector $X = (x, y, z, v_x, v_y, v_z)$ derived from i th TLE set to the time point $t_{j,k}$;
- $X(t_{j,k})$ the predicted state vector derived from j th TLE set itself to the time point $t_{j,k}$, $X(t_{j,k}) \equiv X(t_j/t_{j,k})$;
- $\Delta X(t_i/t_{j,k})$ the residual vector derived from the differencing between $X(t_i/t_{j,k})$ and $X(t_{j,k})$, $\Delta X(t_i/t_{j,k}) = X(t_i/t_{j,k}) - X(t_{j,k})$.

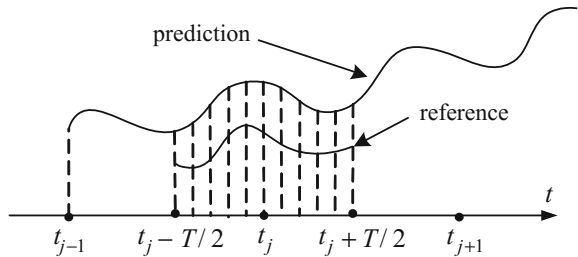
Take an orbital period-long time interval centered at the j th TLE's epoch t_j , as illustrated in Fig. 3.2. This time interval is then separated into n_{seg} segments so we get $(n_{\text{seg}} + 1)$ time points.

$$t_{j,k} = t_j - \frac{T}{2} + k \frac{T}{n}, \quad k = 0, 1, 2, \dots, n_{\text{seg}} \quad (3.2)$$

Figure 3.2 illustrates a case where n_{seg} equals to ten. At each time point $t_{j,k}$, predicted state vector $X(t_i/t_{j,k})$ derived from i th ($i < j$) TLE and state vector $X(t_{j,k})$ derived from j th TLE set itself can be obtained. Differencing $X(t_i/t_{j,k})$ and $X(t_{j,k})$ we can get the residual vector $\Delta X(t_i/t_{j,k})$ and corresponding temporal difference $\Delta t_{i/j,k}$ and the mean anomaly $M_{j,k}$ at time points $t_{j,k}$.

The i th TLE then will be propagated to all subsequent TLE's orbital period-long time interval and compared with state vector derived from subsequent TLE to get residual data as illustrated in Fig. 3.3 until temporal difference exceeds time threshold t_T . So the i th TLE's prediction residual data and corresponding temporal

Fig. 3.2 Generation of residuals between predicted and reference state in whole period



differences and mean anomalies can be obtained. Every residual vector is associated with a time parameter and an angular parameter.

Repeating the process for all available TLEs, we can obtain a series of residual vectors and corresponding temporal differences and mean anomalies.

$$\{\Delta X_1(\Delta t_1, M_1), \Delta X_2(\Delta t_2, M_2), \dots, \Delta X_N(\Delta t_N, M_N)\} \tag{3.3}$$

All the state residual vectors are transformed from Earth Centered Inertial (ECI) coordinate system to satellite-based UNW (in-track, normal, and cross-track) coordinate system, as

$$\Delta X = [\Delta U, \Delta N, \Delta W, \Delta V_U, \Delta V_N, \Delta V_W] \tag{3.4}$$

Therefore a $(N \times 8)$ residual data matrix could be established as Eq. (3.5), the first column of the matrix is temporal difference Δt which will not exceed the time threshold t_T , the second column is mean anomaly $M \in [-\pi, +\pi]$, and the third to eighth columns are six components of residual state vector in UNW coordinate system.

$$\begin{bmatrix} \Delta t_1 & M_1 & \Delta U_1 & \Delta N_1 & \Delta W_1 & \Delta V_{U1} & \Delta V_{N1} & \Delta V_{W1} \\ \Delta t_2 & M_2 & \Delta U_2 & \Delta N_2 & \Delta W_2 & \Delta V_{U2} & \Delta V_{N2} & \Delta V_{W2} \\ \Delta t_3 & M_3 & \Delta U_3 & \Delta N_3 & \Delta W_3 & \Delta V_{U3} & \Delta V_{N3} & \Delta V_{W3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Delta t_N & M_N & \Delta U_N & \Delta N_N & \Delta W_N & \Delta V_{UN} & \Delta V_{NN} & \Delta V_{WN} \end{bmatrix} \begin{matrix} \Delta t \in [0, +\infty) \\ M \in [-\pi, +\pi] \end{matrix} \tag{3.5}$$

So far, we have gotten a residual data sample matrix, this data matrix is the basic for the error analysis; all the analyses below are based on this data matrix. Let the temporal difference threshold $t_T = 3$ days and segment number $n = 20$, Figs. C.1–

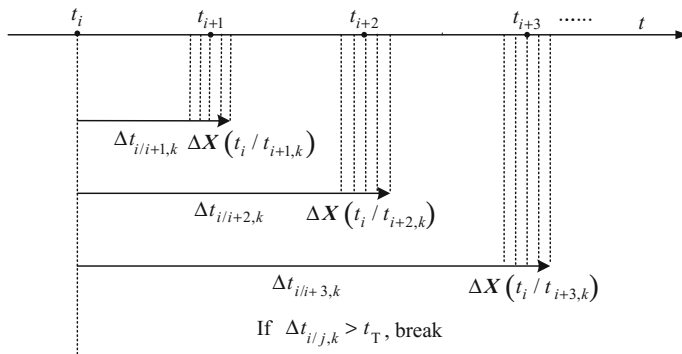


Fig. 3.3 Generation of i th TLE’s prediction residual to subsequent TLEs

C.8 in Appendix C illustrate the TLE's prediction residual state components in UNW coordinate system versus temporal difference and versus mean anomaly of four selected objects, respectively. It is evidently that there are periodic characteristics in all six components. Figures C.5 and C.6 show that errors near the perigee where mean anomalies are about zeros are larger than errors near the apogee where mean anomalies are about $\pm 180^\circ$. There is only one time parameter in polynomial fitting function so that cannot reflect the periodic variation of error in an orbital period.

3.3 Residual Data Preprocessing

Before further process, it is necessary to preprocess the residuals data. The pre-processing of residual data consists of data binning and outlier detection.

3.3.1 Residual Data's 2D Binning

Osweiler [10] binned the residual data according to the corresponding epoch time difference Δt without consideration of the regular and discrete characteristics of TLE's epoch. It was supposed that TLEs are provided at unequal time interval therefore the epoch time differences resulting from the pair-wise comparison can be any value. Bin intervals of one day were chosen to keep the overall computational time reasonable for his study.

The residual data in this study has two parameters, Fig. 3.4a illustrates the temporal differences and mean anomalies of residual data for LEO object, so two-dimensional data binning must be done. Because of the regular and discrete characteristics of TLE's epoch, the epoch time differences will concentrate on a series of discrete values, and these values equal to period or multiple of period. Figure 3.5 demonstrates this characteristic clearly. Considering the discrete characteristics, it is convenient to bin the time parameters according to the ratio of temporal difference Δt to orbital period T . Generally, the ratios $n_{\Delta t/T} = \Delta t/T$ are not always integers, so we round it to the nearest integer $N_{\Delta t/T}$. Then the time parameters can be binned according to number $N_{\Delta t/T}$.

As described before, an orbital period has been separated into n segments, so each segment's length is $M_{\text{bin}} = 360/n$. It is convenient to bin the angular parameters according to the ratio of mean anomaly M to segment's length M_{bin} . Similarly, the ratios $n_M = M/M_{\text{bin}}$ are not always integers; they are rounded to the nearest integer N_M . Then the angular parameters can be binned according to number N_M . Thus, two associated integral parameters are obtained for every residual data, residual data can be binned two-dimensionally according to these two parameters.

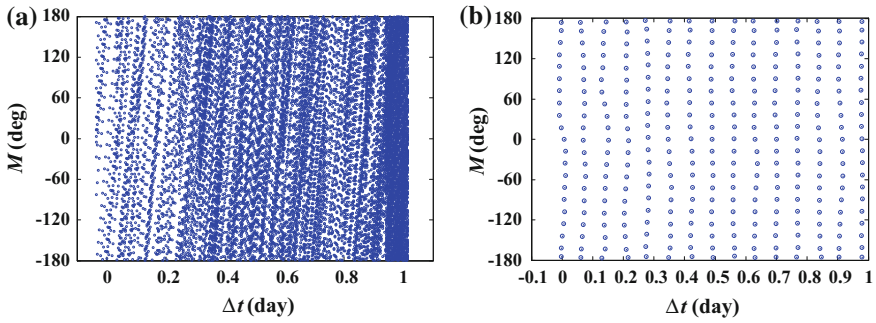


Fig. 3.4 Residual data two-dimensional binning (a left, before binning; b right, after binning)

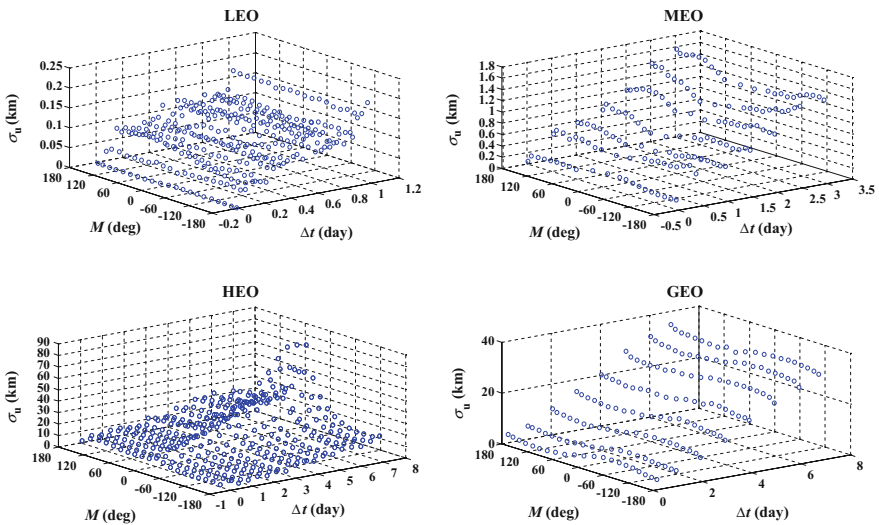


Fig. 3.5 In-track (U) position error standard deviations for each data bins of four objects

Figure 3.4b illustrates the scattered distribution of every data bins’ time and angular parameters.

If the threshold of propagation duration is t_T , orbital period is T , the minimum and maximum value of $N_{\Delta t/T}$ will be 0 and $\text{int}(t_T/T)$, the minimum and maximum value of N_M will be 0 and n_{seg} . The total number of data bins will be

$$N_{\text{bin}} = \left(\text{int}\left(\frac{t_T}{T}\right) + 1 \right) (n_{\text{seg}} + 1) \tag{3.6}$$

The number of data bins in Fig. 3.4 is $N_{\text{bin}} = 15 \times 21 = 315$.

3.3.2 Outlier Detection of Residual Data

Due to orbit maneuver, TLE generating error, or other reasons, there are inevitable outliers in TLE sets. Consequently, there are outliers in residual data. After data binning, outliers in each data bins must be detected. A Mahalanobis-distance-based outlier detecting method is used [19], if Mahalanobis distance d_M from a data point to the distributing center of the bin exceeds distance threshold, then this data point will be seen as an outlier and eliminated.

After data binning and outlier detection, a series of data bins which include a set of residual data having similar temporal differences and mean anomalies are obtained. We could calculate mean temporal difference t_k , means of mean anomaly M_k , and position error standard deviations of in-track, normal, and cross-track components σ_{Uk} , σ_{Nk} , σ_{Wk} of each data bin.

Figure 3.5 shows the in-track standard deviations for each data bins of four objects. We also can see the temporal evolution and periodic characteristics of prediction error from Fig. 3.5. The periodic characteristics are especially remarkable for the HEO object.

3.4 Form of Fitting Function Considering Periodicity

The determination process of error fitting function can be divided into two steps: first determining the form of fitting function, then the coefficients of fitting function.

3.4.1 Error of Orbital Elements and Positions

The relationship between the error of orbital elements and the R , S , W components of position error has been provided in Sect. 2.4.2 of Chap. 2. Substitution of the expression of δa and δM into the relationship yields

$$\left\{ \begin{array}{l} \delta R = \frac{r}{a} \delta a_0 + \frac{r}{a} \delta \dot{a} \cdot \tau - a \delta e \cos f + \frac{ae \delta M_0}{\sqrt{1-e^2}} \sin f - \frac{3ne \delta a_0}{2\sqrt{1-e^2}} \tau \sin f \\ \quad - \frac{3ne \delta \dot{a}}{4\sqrt{1-e^2}} \tau^2 \sin f \\ \quad = C_{R1} + C_{R2} \tau + C_{R3} \cos f + C_{R4} \sin f + C_{R5} \tau \sin f + C_{R6} \tau^2 \sin f \\ \delta S = r(\delta \omega + \cos i \delta \Omega) + \frac{a^2 \sqrt{1-e^2}}{r} \delta M_0 - \frac{3na \delta a_0 \sqrt{1-e^2}}{2r} \tau \\ \quad - \frac{3na \delta \dot{a} \sqrt{1-e^2}}{4r} \tau^2 + \frac{2r \delta e}{1-e^2} \sin f + \frac{re \delta e}{2(1-e^2)} \sin 2f \\ \quad = C_{S1} + C_{S2} \tau + C_{S3} \tau^2 + C_{S4} \sin f + C_{S5} \sin 2f \\ \delta W = r(\sin \omega \delta i - \sin i \cos \omega \delta \Omega) \cos f + r(\cos \omega \delta i + \sin i \sin \omega \delta \Omega) \sin f \\ \quad = C_{W1} \cos f + C_{W2} \sin f \end{array} \right. \quad (3.7)$$

Equation (3.7) indicates that the R , S , W components of position error can be commonly written as a combination of the polynomial with respect to the prediction time τ and the trigonometric function series with respect to the true anomaly f . Set the order of the time polynomial to be $n = 2$, and the order of the trigonometric function series to be $m = 2$, the errors of three directions can be written as the universal form

$$\begin{aligned} \delta &= A_{0,0} + A_{0,1} \cos f + B_{0,1} \sin f + A_{0,2} \cos 2f + B_{0,2} \sin 2f \\ &\quad + A_{1,0}\tau + A_{1,1}\tau \cos f + B_{1,1}\tau \sin f + A_{1,2}\tau \cos 2f + B_{1,2}\tau \sin 2f \\ &\quad + A_{2,0}\tau^2 + A_{2,1}\tau^2 \cos f + B_{2,1}\tau^2 \sin f + A_{2,2}\tau^2 \cos 2f + B_{2,2}\tau^2 \sin 2f \quad (3.8) \\ &= \sum_{i=0}^2 \sum_{j=0}^2 \tau^i (A_{i,j} \cos jf + B_{i,j} \sin jf) \end{aligned}$$

Equation (3.8) is essentially the expression of Poisson series which will be discussed below.

3.4.2 Error Propagation Results of C-W Equations

Chapter 2 presented the covariance's propagation formulae of near-circular orbit based on the C-W Equations. The diagonal terms of covariance matrix can be written as the polynomial and trigonometric form:

$$\left\{ \begin{aligned} \sigma_x^2 &= 4 \left(4\sigma_{x_0}^2 + \frac{\sigma_{v_{y0}}^2}{n^2} \right) - 8 \left(3\sigma_{x_0}^2 + \frac{\sigma_{v_{y0}}^2}{n^2} \right) \cos(nt) + \left(9\sigma_{x_0}^2 + \frac{4\sigma_{v_{y0}}^2}{n^2} \right) \cos^2(nt) \\ &\quad + \frac{\sigma_{v_{y0}}^2}{n^2} \sin^2(nt) \\ \sigma_y^2 &= \left(\sigma_{y_0}^2 + \frac{4}{n^2} \sigma_{v_{x0}}^2 \right) + 3 \left(3\sigma_{v_{y0}}^2 + n^2 \sigma_{x_0}^2 \right) t^2 - 6 \left(n\sigma_{x_0}^2 + \frac{4}{n} \sigma_{v_{y0}}^2 \right) t \sin(nt) \\ &\quad - \frac{8\sigma_{v_{x0}}^2}{n^2} \cos(nt) + \frac{4\sigma_{v_{x0}}^2}{n^2} \cos^2(nt) + \left(3\sigma_{x_0}^2 + \frac{16}{n^2} \sigma_{v_{y0}}^2 \right) \sin^2(nt) \\ \sigma_z^2 &= \sigma_{z_0}^2 \cos^2(nt) + \frac{\sigma_{v_{z0}}^2}{n^2} \sin^2(nt) \\ \sigma_{v_x}^2 &= \left(n^2 \sigma_{y_0}^2 + 4\sigma_{v_{x0}}^2 \right) + \left(36n^4 \sigma_{x_0}^2 + 9n^2 \sigma_{v_{y0}}^2 \right) t^2 - \left(36n^3 \sigma_{x_0}^2 + 12n \sigma_{v_{y0}}^2 \right) t \sin(nt) \\ &\quad - 4\sigma_{v_{x0}}^2 \cos(nt) + \sigma_{v_{x0}}^2 \cos^2(nt) + \left(9n^2 \sigma_{x_0}^2 + 4\sigma_{v_{y0}}^2 \right) \sin^2(nt) \\ \sigma_{v_y}^2 &= \left(4n^2 \sigma_{x_0}^2 + \sigma_{v_{y0}}^2 \right) - \left(12n^2 \sigma_{x_0}^2 + 4\sigma_{v_{y0}}^2 \right) \cos(nt) \\ &\quad + \left(9n^2 \sigma_{x_0}^2 + 4\sigma_{v_{y0}}^2 \right) \cos^2(nt) + \sigma_{v_{y0}}^2 \sin^2(nt) \\ \sigma_{v_z}^2 &= \sigma_{v_{z0}}^2 \cos^2(nt) + n^2 \sigma_{z_0}^2 \sin^2(nt) \end{aligned} \right. \quad (3.9)$$

For convenience let ω denote the angular velocity. By introducing coefficients the variances can be expressed as the general form:

$$\sigma^2(t) = c_0 + c_1t + c_2t^2 + c_3 \cos \omega t + c_4 \sin \omega t + c_5 \cos^2 \omega t + c_6 \sin^2 \omega t + c_7t \cos \omega t + c_8t \sin \omega t \quad (3.10)$$

where the quadratic terms of trigonometric function can be transformed to the high-frequency terms

$$\sigma^2(t) = c'_0 + c'_1t + c'_2t^2 + c'_3 \cos \omega t + c'_4 \sin \omega t + c'_5 \cos 2\omega t + c'_6 \sin 2\omega t + c'_7t \cos \omega t + c'_8t \sin \omega t \quad (3.11)$$

Equation (3.11) reveals that the propagation function of error variance consists of three parts:

- (1) Polynomial term ($c'_0 + c'_1t + c'_2t^2$), describing the secular evolution of prediction error.
- (2) Trigonometric terms ($\cos \omega t, \sin \omega t, \cos 2\omega t, \sin 2\omega t$), describing the periodic characteristics of prediction error whose amplitude keeps constant.
- (3) Mixed terms ($t \cos \omega t, t \sin \omega t$), describing the periodic term whose amplitude grow with time.

The coefficients of these terms are determined by the angular velocity ω of orbit and initial error variances. For the error propagation described by the C-W Equation, the S direction position error and the R direction velocity error have the polynomial terms and mixed terms with the secular evolution, and the order of polynomial term is two; error variances in other directions only have constant terms and trigonometric terms, and the order of trigonometric function is two.

3.4.3 Poisson Series

The previous analysis reveals that the propagation functions of both orbital prediction error and error variance are combination of polynomial function and trigonometric function, which is essentially a special series—Poisson series.

1. General Form of Poisson Series

Poisson series appears frequently in celestial mechanics and more generally in perturbation theories for non-linear mechanics or non-linear differential equations [20–30]. Poisson series also called polynomial-trigonometric series, has the general form as [20]:

$$P = \sum_{i_1, \dots, i_n} \sum_{j_1, \dots, j_m} C_{i_1, \dots, i_n}^{j_1, \dots, j_m} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n} \begin{pmatrix} \sin \\ \cos \end{pmatrix} (j_1 \phi_1 + \dots + j_m \phi_m) \tag{3.12}$$

where x_1, x_2, \dots, x_n are polynomial variables, $\phi_1, \phi_2, \dots, \phi_m$ are trigonometric variables, respectively. Coefficients C_i^j may be represented as rational, floating-point or complex numbers. The variables x_1, x_2, \dots, x_n and $\phi_1, \phi_2, \dots, \phi_m$ are rational numbers, the indices i_1, i_2, \dots, i_n and j_1, j_2, \dots, j_m are integer numbers. The summation is performed over all integer values of indices i_1 to i_n and j_1 to j_m .

Poisson series have several important properties [20]. The results of addition, subtraction, multiplication and differentiation (with respect to any of the variables) of Poisson series are again Poisson series. This property allows us to fit either standard deviation or variance of prediction error. Considering convenience for using, this section handles the standard deviation. If the polynomial arguments are raised to non-negative powers, this also holds for the result of substituting one Poisson series for one of the polynomial arguments of another, and for the case of integration with respect to a polynomial argument. Integration with respect to a trigonometric argument also results in a Poisson series if the original series contains no terms which are constant with respect to that argument (i.e., the ‘periodic part’ of the series). Finally, a Poisson series is obtained if a substitution of one series for an angular variable in another series is made by means of an expansion in (truncated) power series.

2. (1, 1)-Style Poisson Series with One Polynomial Variable and One Trigonometric Variable

There is one polynomial variable (temporal difference) and one trigonometric variable (mean anomaly) in the residual data, hence Eq. (3.12) could be simplified to (1, 1)-style Poisson series

$$P(t, \phi) = \sum_{i=0}^n \sum_{j=0}^m t^i (A_{i,j} \cos j\phi + B_{i,j} \sin j\phi) \tag{3.13}$$

where t and ϕ are polynomial variables and trigonometric variables, $A_{i,j}$ and $B_{i,j}$ are Poisson coefficients, n and m are the maximum orders of polynomial terms and trigonometric terms, respectively. Because $\sin j\phi = 0$ when $j = 0$, coefficients $B_{i,0} (i = 0, \dots, n)$ are meaningless, number of coefficients is $(n + 1)(2m + 1)$.

Equation (3.13) can be expanded to

$$\begin{aligned}
P(t, \phi) = & A_{0,0} + A_{0,1} \cos \phi + A_{0,2} \cos 2\phi + \cdots + A_{0,m} \cos m\phi + B_{0,1} \sin \phi \\
& + B_{0,2} \sin 2\phi + \cdots + B_{0,m} \sin m\phi \\
& + A_{1,0}t + A_{1,1}t \cos \phi + A_{1,2}t \cos 2\phi + \cdots + A_{1,m}t \cos m\phi \\
& + B_{1,1}t \sin \phi + B_{1,2}t \sin 2\phi + \cdots + B_{1,m}t \sin m\phi \\
& + A_{2,0}t^2 + A_{2,1}t^2 \cos \phi + A_{2,2}t^2 \cos 2\phi + \cdots + A_{2,m}t^2 \cos m\phi \\
& + B_{2,1}t^2 \sin \phi + B_{2,2}t^2 \sin 2\phi + \cdots + B_{2,m}t^2 \sin m\phi \\
& + \cdots + A_{n,0}t^n + A_{n,1}t^n \cos \phi + A_{n,2}t^n \cos 2\phi + \cdots + A_{n,m}t^n \cos m\phi \\
& + B_{n,1}t^n \sin \phi + B_{n,2}t^n \sin 2\phi + \cdots + B_{n,m}t^n \sin m\phi
\end{aligned} \tag{3.14}$$

Equation (3.14) indicates that expanded form of (1,1)-style Poisson series with one polynomial variable and one trigonometric variable consists of three parts:

- (1) Polynomial terms, the terms with coefficients $A_{i,0}(i = 0, \dots, n)$;
- (2) Trigonometric terms, the terms with coefficients $A_{0,j}(j = 1, \dots, m)$ and $B_{0,j}(j = 1, \dots, m)$;
- (3) Mixed terms, the terms in Eq. (3.14) with coefficients $A_{i,j}(i = 1, \dots, n, j = 1, \dots, m)$ and $B_{i,j}(i = 1, \dots, n, j = 1, \dots, m)$.

The Poisson series can also be written as vector and matrix form

$$\left\{ \begin{array}{l} P(t, \phi) = [1 \quad t \quad t^2 \quad \cdots \quad t^n] \\ \left[\begin{array}{cccc} A_{0,0} & A_{0,1} & \cdots & A_{0,m} \\ A_{1,0} & A_{1,1} & \cdots & A_{1,m} \\ \vdots & \vdots & & \vdots \\ A_{n,0} & A_{n,1} & \cdots & A_{n,m} \end{array} \right] \left[\begin{array}{c} 1 \\ \cos \phi \\ \vdots \\ \cos m\phi \end{array} \right] + \left[\begin{array}{cccc} 0 & B_{0,1} & \cdots & B_{0,m} \\ 0 & B_{1,1} & \cdots & B_{1,m} \\ \vdots & \vdots & & \vdots \\ 0 & B_{n,1} & \cdots & B_{n,m} \end{array} \right] \left[\begin{array}{c} 0 \\ \sin \phi \\ \vdots \\ \sin m\phi \end{array} \right] \end{array} \right\}. \tag{3.15}$$

Where X and Y are polynomial variable vectors, Z is trigonometric variable vectors, A and B are Poisson coefficient matrices (PCM).

$$\left\{ \begin{array}{l} X(t) = (1 \quad t \quad t^2 \quad \cdots \quad t^n)^T \\ Y(\phi) = (1 \quad \cos \phi \quad \cos 2\phi \quad \cdots \quad \cos m\phi)^T \\ Z(\phi) = (0 \quad \sin \phi \quad \sin 2\phi \quad \cdots \quad \sin m\phi)^T \\ A = \left[\begin{array}{cccc} A_{0,0} & A_{0,1} & \cdots & A_{0,m} \\ A_{1,0} & A_{1,1} & \cdots & A_{1,m} \\ \vdots & \vdots & & \vdots \\ A_{n,0} & A_{n,1} & \cdots & A_{n,m} \end{array} \right], B = \left[\begin{array}{cccc} 0 & B_{0,1} & \cdots & B_{0,m} \\ 0 & B_{1,1} & \cdots & B_{1,m} \\ \vdots & \vdots & & \vdots \\ 0 & B_{n,1} & \cdots & B_{n,m} \end{array} \right] \end{array} \right\} \tag{3.16}$$

So the vector and matrix form of (1, 1)-style Poisson is

$$P(t, \phi) = \mathbf{X}^T(t)[\mathbf{A}\mathbf{Y}(\phi) + \mathbf{B}\mathbf{Z}(\phi)] \tag{3.17}$$

In addition, Poisson series can also be expressed as complex form

$$\begin{aligned}
 P(t, \phi) &= \mathbf{X}^T(t) \cdot \mathbf{C} \cdot \mathbf{U}(\phi) \\
 \mathbf{U}(\phi) &= \{e^{ij\phi}\}_{j=-m}^m = (e^{-im\phi} \quad e^{-i(m-1)\phi} \quad \dots \quad 1 \quad \dots \quad e^{i(m-1)\phi} \quad e^{im\phi})^T \\
 \mathbf{C} &= \begin{bmatrix} C_{0,-m} & \dots & C_{0,-1} & C_{0,0} & C_{0,1} & \dots & C_{0,m} \\ C_{1,-m} & \dots & C_{1,-1} & C_{1,0} & C_{1,1} & \dots & C_{1,m} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ C_{n,-m} & \dots & C_{n,-1} & C_{n,0} & C_{n,1} & \dots & C_{n,m} \end{bmatrix}_{(n+1) \times (2m+1)}
 \end{aligned} \tag{3.18}$$

where $C_{i,j}$ are conjugate complex coefficients, $C_{k,j} = C_{k,-j}^*$, so the number of independent coefficients is still $(n + 1)(2m + 1)$.

3. Square of Poisson Series

Poisson series have several important properties; one of them is that the results of addition, subtraction, multiplication and differentiation (with respect to any of the variables) of Poisson series are again Poisson series.

To validate the property, we will obtain the formulae of the square of Poisson series $P^2(t, \phi)$ based on the complex form of Poisson series. Equation (3.18) implies that the expression of $P(t, \phi)$ is essentially a sum of a series of terms, each term is the product of power of t and exponential of $i\phi$. The general form of the sum of a series of terms can be expressed as

$$S = x_1 + x_2 + \dots + x_N = \sum_{i=1}^N x_i \tag{3.19}$$

The square of S is

$$\begin{aligned}
 S^2 &= (x_1 + x_2 + \dots + x_N)(x_1 + x_2 + \dots + x_N) = \left(\sum_{i=1}^N x_i\right) \left(\sum_{i=1}^N x_i\right) \\
 &= \sum_{i=1}^N \sum_{j=1}^N x_i x_j
 \end{aligned} \tag{3.20}$$

Equation (3.20) reveals the square of sum is also sum, and the terms in new sum are the product of two terms which come from original sum. The product of any two terms from Eq. (3.18) is

$$(C_{k_1 j_1} t^{k_1} e^{ij_1 \phi}) \cdot (C_{k_2 j_2} t^{k_2} e^{ij_2 \phi}) = C_{k_1 j_1} C_{k_2 j_2} t^{k_1 + k_2} e^{i(j_1 + j_2) \phi} \quad (3.21)$$

The product of any two terms from Poisson series is still the product of power of t and exponential of $i\phi$, just has different maximum order. The k_1 and k_2 in Eq. (3.21) satisfy $0 \leq k_1 \leq n$ and $0 \leq k_2 \leq n$, so $k_1 + k_2$ satisfy $0 \leq k_1 + k_2 \leq 2n$; j_1 and j_2 satisfy $-m \leq j_1 \leq m$ and $-m \leq j_2 \leq m$, so $j_1 + j_2$ satisfy $-2m \leq j_1 + j_2 \leq 2m$. Therefore, the square $P^2(t, \phi)$ of a Poisson series $P(t, \phi)$ is still a Poisson series, with the maximum order of polynomial terms and trigonometric terms are $2n$ and $2m$ as follow

$$P^2(t, \phi) = \sum_{k=0}^{2n} \sum_{j=-2m}^{2m} \tilde{C}_{k,j} t^k e^{ij\phi} \quad (3.22)$$

The fact that the square of Poisson series are again Poisson series implies either standard deviation or variance of prediction error can be fitted by Poisson series, just with different orders and coefficient matrices. Considering convenience for using, this section handles the standard deviation.

3.5 Error Coefficient Matrix Fitting

3.5.1 Least Squares Fitting of Poisson Series

To describe temporal evolution and periodic characteristics of orbital prediction error by Poisson series, we fit Poisson series over the full available data bins using least squares method. We have three-dimensional discrete data set $\{t_k, \phi_k, S_k\}_{k=1}^N$ for a certain error component. Least squares method finds the Poisson coefficient matrix (PCM) \mathbf{A} and \mathbf{B} , make the objective function Eq. (3.23) to be minimum.

$$F = \sum_{k=1}^N [\sigma_k - p(t_k, \phi_k)]^2 \quad (3.23)$$

Each data set is substituted into a linear equation about unknown coefficients:

$$\begin{aligned} \sigma_1 &= A_{0,0} + A_{0,1} \cos \phi_1 + \cdots + A_{0,m} \cos m\phi_1 + B_{0,1} \sin \phi_1 + \cdots + B_{0,m} \sin m\phi_1 + \cdots + A_{n,0} t_1^n \\ &\quad + A_{n,1} t_1^n \cos \phi_1 + \cdots + A_{n,m} t_1^n \cos m\phi_1 + B_{n,1} t_1^n \sin \phi_1 + \cdots + B_{n,m} t_1^n \sin m\phi_1 + \varepsilon_1 \\ \sigma_2 &= A_{0,0} + A_{0,1} \cos \phi_2 + \cdots + A_{0,m} \cos m\phi_2 + B_{0,1} \sin \phi_2 + \cdots + B_{0,m} \sin m\phi_2 + \cdots + A_{n,0} t_2^n \\ &\quad + A_{n,1} t_2^n \cos \phi_2 + \cdots + A_{n,m} t_2^n \cos m\phi_2 + B_{n,1} t_2^n \sin \phi_2 + \cdots + B_{n,m} t_2^n \sin m\phi_2 + \varepsilon_2 \\ &\dots \\ \sigma_N &= A_{0,0} + A_{0,1} \cos \phi_N + \cdots + A_{0,m} \cos m\phi_N + B_{0,1} \sin \phi_N + \cdots + B_{0,m} \sin m\phi_N + \cdots + A_{n,0} t_N^n \\ &\quad + A_{n,1} t_N^n \cos \phi_N + \cdots + A_{n,m} t_N^n \cos m\phi_N + B_{n,1} t_N^n \sin \phi_N + \cdots + B_{n,m} t_N^n \sin m\phi_N + \varepsilon_3 \end{aligned} \quad (3.24)$$

Let

$$\begin{aligned}
 \mathbf{S} &= (\sigma_1 \ \sigma_2 \ \cdots \ \sigma_N)_{N \times 1}^T, \quad \boldsymbol{\varepsilon} = (\varepsilon_1 \ \varepsilon_2 \ \cdots \ \varepsilon_N)_{N \times 1}^T \\
 \boldsymbol{\theta} &= \left(A_{0,0} \ A_{0,1} \ \cdots \ A_{0,m} \ B_{0,1} \ \cdots \ B_{0,m}; \ \cdots \ ; A_{n,0} \ A_{n,1} \ \cdots \ A_{n,m} \ B_{n,1} \ \cdots \ B_{n,m} \right)_{(n+1)(2m+1) \times 1}^T \\
 \mathbf{H} &= \begin{bmatrix} 1 & \cos \phi_1 & \cdots & \cos m\phi_1 & \sin \phi_1 & \cdots & \sin m\phi_1; & \cdots & ; t_1^n & t_1^n \cos \phi_1 & \cdots & t_1^n \cos m\phi_1 & t_1^n \sin \phi_1 & \cdots & t_1^n \sin m\phi_1 \\ 1 & \cos \phi_2 & \cdots & \cos m\phi_2 & \sin \phi_2 & \cdots & \sin m\phi_2; & \cdots & ; t_2^n & t_2^n \cos \phi_2 & \cdots & t_2^n \cos m\phi_2 & t_2^n \sin \phi_2 & \cdots & t_2^n \sin m\phi_2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & \cos \phi_N & \cdots & \cos m\phi_N & \sin \phi_N & \cdots & \sin m\phi_N; & \cdots & ; t_N^n & t_N^n \cos \phi_N & \cdots & t_N^n \cos m\phi_N & t_N^n \sin \phi_N & \cdots & t_N^n \sin m\phi_N \end{bmatrix}_{N \times (n+1)(2m+1)}
 \end{aligned} \tag{3.25}$$

then

$$\mathbf{S}_{N \times 1} = \mathbf{H}_{N \times (n+1)(2m+1)} \boldsymbol{\theta}_{(n+1)(2m+1) \times 1} + \boldsymbol{\varepsilon}_{N \times 1} \tag{3.26}$$

the least squares estimation $\hat{\boldsymbol{\theta}}$ of coefficient vector $\boldsymbol{\theta}$ satisfy the canonical equation:

$$\mathbf{H}^T \mathbf{H} \hat{\boldsymbol{\theta}} = \mathbf{H}^T \mathbf{S} \tag{3.27}$$

This equation is a $(n+1)(2m+1)$ th order linear equation set. When \mathbf{H} is column non-singular, i.e.

$$\text{rank}(\mathbf{H}) = (n+1)(2m+1) < N \tag{3.28}$$

the matrix $\mathbf{H}^T \mathbf{H}$ is non-singular, $\text{rank}(\mathbf{H}^T \mathbf{H}) = (n+1)(2m+1)$, the inverse matrix of $\mathbf{H}^T \mathbf{H}$ exists, then

$$\hat{\boldsymbol{\theta}} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{S} \tag{3.29}$$

When the estimation of coefficient vector $\hat{\boldsymbol{\theta}}$ is obtained, the Poisson coefficient matrices (PCM) \mathbf{A} and \mathbf{B} can easily be got from Eq. (3.25).

3.5.2 Determination of Poisson Series' Order

Least squares method requires that the number of coefficients to be fitted is less than the number of data sets, which means in this work the maximum orders of Poisson series should satisfy the in equation Eq. (3.28). In Eq. (3.28) n and m are the maximum orders of polynomial terms and trigonometric terms, N is the number of data bins. Substituting Eq. (3.6) into Eq. (3.28), we can obtain

$$(n + 1)(2m + 1) < \left(\left\lceil \frac{t_T}{T} \right\rceil + 1 \right) (n_{\text{seg}} + 1) \quad (3.30)$$

From Eq. (3.30) we know if n and m satisfy the following inequations the Poisson coefficient matrix could be fitted using least squares method.

$$n < \left\lceil \frac{t_T}{T} \right\rceil, \quad m < \frac{n_{\text{seg}}}{2} \quad (3.31)$$

The dimension number of unknown coefficients in least squares method is $(n + 1)(2m + 1)$. If $n = 2$ and $m = 2$ the dimension number is 15. The dimension of coefficients and the order of coefficient matrix increase with increase of the maximum order of Poisson series, which is computationally burdensome and leads to oversized coefficient matrix. It is recommended that $n \leq 2$ and $m \leq 4$ in practice.

The algorithms implemented in this research are coded and performed in MATLAB.

3.6 Results and Analysis

3.6.1 Fitting Result of Position Error

With the methodology developed, the residual data can be generated and preprocessed, the Poisson coefficient matrices of each error component can be fitted using least squares method. Some results will be given in this section.

The propagation duration thresholds t_T are one and three days for LEO and MEO objects, and seven days for HEO and GEO objects, respectively. The number of segment $n_{\text{seg}} = 20$. The maximum orders of polynomial term and trigonometric term are $n = 2$ and $m = 4$, respectively. Therefore, the Poisson coefficient matrices are two 3×5 matrices. The Poisson coefficient matrices of in-track (U) error standard deviation for four objects are given in Table 3.2.

Figures C.9–C.12 in Appendix C illustrate the in-track (U), normal (N), and cross-track (W) error standard deviation's Poisson series fitting curved surfaces of four objects. It is noticeable that the periodic characteristics exist in prediction error of all three directions of four objects more or less, and HEO and MEO orbit's periodic characteristics are more obvious, especially the HEO orbit.

Table 3.2 Poisson coefficient matrix of in-track (U) error standard deviation for four objects

LEO Object					
A_U	0.0311615712	-0.0029029078	0.0005102982	0.0004477086	0.0000482616
	0.1124245624	0.0067061297	-0.0030056393	-0.0033477243	-0.0006128871
	0.0209600339	-0.0026381835	0.0007039691	0.0020221898	0.0001487366
B_U	0	0.0038191894	0.0022722895	-0.0000696655	0.0000138184
	0	-0.0020553761	-0.0040007188	-0.0015870270	0.0016054081
	0	-0.0006137858	0.0010904506	0.0009060865	-0.0005683997
MEO Object					
A_U	0.2418092045	-0.0026798490	-0.0120973106	-0.0063728817	-0.0119340684
	0.3416823133	-0.0847030456	-0.0289816633	-0.0064566465	0.0124470628
	0.0179197673	0.0108429227	0.0067345439	0.0009171846	-0.0042677763
B_U	0	0.0978250153	0.0373950962	-0.0009417343	0.0125371224
	0	-0.0255268411	-0.0227488696	0.0083173229	0.0032436791
	0	0.0170423174	0.0072646921	-0.0022797763	-0.0023141290
HEO Object					
A_U	3.4240982423	-0.7535455752	0.5830053208	0.5296399607	0.2080734297
	3.0063481052	2.3755155457	0.9120809048	0.3738397827	0.2566062423
	0.2344403167	0.1147006916	0.0987109446	0.0711921664	0.0525966451
B_U	0	0.5444392209	-0.0081093083	0.0409812162	0.0890144982
	0	-0.1420232095	-0.0249054348	-0.0118579123	-0.0427719359
	0	-0.0005894320	0.0015245316	0.0030240224	0.0007457061
GEO Object					
A_U	1.8446800637	0.0951975359	0.1527918715	-0.0066690563	0.0555656114
	4.3571671110	0.0160042137	-0.0654378953	-0.0430340728	-0.0759212314
	0.0720428147	-0.0131970730	0.0065039086	0.0060066348	0.0090739755
B_U	0	0.1893789414	-0.0802648448	-0.0245832745	-0.0171814120
	0	0.1982533435	0.0584842234	0.0315511664	-0.0243514640
	0	-0.0230226125	-0.0056849610	-0.0038554681	0.0005806607

3.6.2 Effect of Poisson Coefficient Matrix

Every part of the Poisson coefficient matrices A and B has specific significations. The elements in first column of matrix A are polynomial coefficients, describing secular evolution of prediction error; the elements in first column of matrix B are defined to be zeros. The elements except first one in first row of both A and B are trigonometric coefficients, describing the periodic characteristics of prediction error. Other elements in A and B are mixed coefficients, describing the periodic term whose amplitude grow with time.

1. Effect of polynomial terms

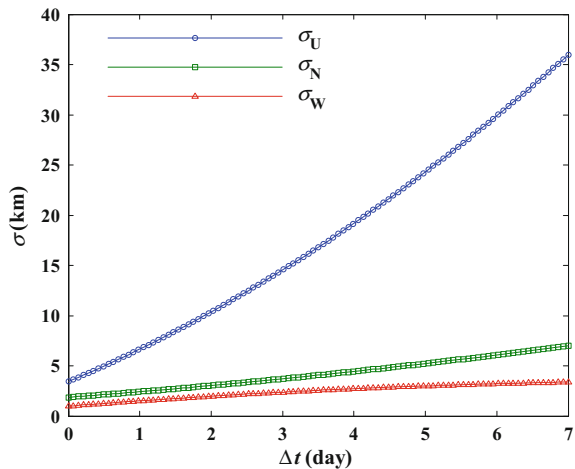
The elements in first column of Poisson coefficient matrix \mathbf{A} are polynomial coefficients of Poisson series, describing the secular evolution of prediction error. If we only take this part of \mathbf{A} into account and set other elements in \mathbf{A} and each element in \mathbf{B} to be zeros, the Poisson series has been degraded to polynomial series. This means the Poisson series fitting method is compatible with polynomial fitting methods developed by Peterson et al. [8]., Deguine et al. [9], and Osweiler [10]. The polynomial fitting is a special case of Poisson series fitting when the trigonometric coefficients are zeros.

The error standard deviation curves of HEO object derived from the degraded Poisson series are shown in Fig. 3.6. In this case only secular evolution of error can be expressed. The normal (N) and cross-track (W) errors are relatively small in magnitude and have near-constant growth rates, the in-track (U) error, however, is much larger and grow more rapidly. The in-track (U) component's standard deviation is one order of magnitude larger than other two components.

2. Effect of trigonometric terms

The elements except first one in first row of both \mathbf{A} and \mathbf{B} are trigonometric coefficients of Poisson series, describing the periodic characteristics of prediction error whose amplitude keeps constant. If we set all the elements in \mathbf{A} and \mathbf{B} except their first row and first column to be zeros, i.e. all the coefficients of mixed terms are zeros, the Poisson series will be degraded to a simple summation of polynomial series and trigonometric series. Figure 3.7 illustrates in-track (U) error standard deviation curved surface of HEO object derived from this kind of Poisson series. Comparing Fig. 3.7 to Fig. 3.6 we see that the trigonometric terms just add periodic variation to polynomial error evolution function.

Fig. 3.6 Error standard deviation curves derived from Poisson series only consists of polynomial terms



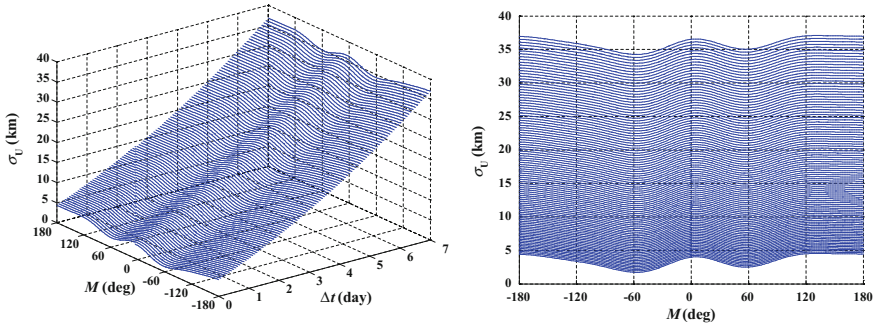


Fig. 3.7 Error standard deviation curve surface derived from Poisson series consists of both polynomial and trigonometric terms

3. Effect of mixed terms

The elements in **A** and **B** except their first row and first column are coefficients of mixed term, describing the periodic term whose amplitude grow with time. The actual characteristic of orbital prediction error is combination of polynomial effect, trigonometric effect, and mixed effect, as illustrated in Figures C.9–C.12. Polynomial, trigonometric, and mixed terms of Poisson series are combined together, describe variation of prediction error with respect to temporal difference and on-orbit position.

4. Significance of Poisson series

In order to illuminate the significance of taking mean anomaly as variable and using Poisson series as fitting function, at temporal difference $\Delta t = 5$ days, Fig. 3.8 is a profile of in-track, normal, and cross-track error standard deviations curved surface of HEO object with respect to mean anomaly.

Fig. 3.8 Profile of in-track, normal, and cross-track error standard deviations curved surface of HEO object with respect to mean anomaly while $\Delta t = 5$ days

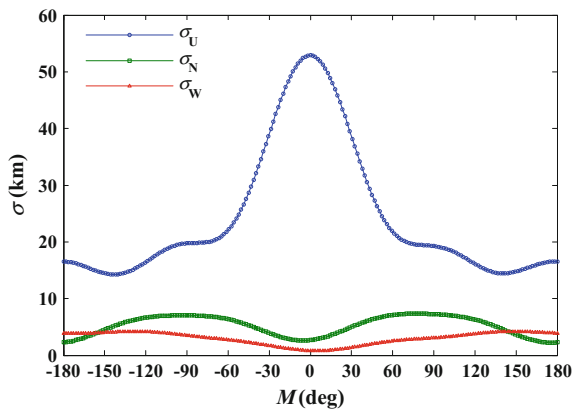


Figure 3.8 shows several properties of error's periodicity. Firstly, orbital error's in-track component is one or two orders of magnitude larger than normal and cross-track components. Secondly, all in-track, normal, and cross-track error have periodic characteristics, but the variation of in-track error in an orbital period is more obvious. Thirdly, in-track error is maximum near the orbital perigee ($M = 0$) and small near the apogee ($M = \pm 180^\circ$), whereas normal and cross-track errors are maximum near $M = \pm 90^\circ$ and small both near the perigee and apogee. In conclusion, it is significant to consider periodic characteristics and introduce angular variable to the evaluation of orbital error.

This work can be further improved in some aspects. For example, it is too rough to take predicted states at or near the epoch as references. TLE sets are the best fit to observations within the fitting interval. This does not imply that the predicted states at epoch are accurate, and the errors at epoch are not necessarily the minimum. One possible improvement is to take states resulting from an orbit determination using pseudo-observations derived from the TLE data followed by a numerical propagation as reference states. Flohrer et al. [6], Levit and Marshall [16] have provided the orbit fitting method to pseudo-observations derived from successive TLEs using high-precision numerical orbital model and batch least squares differential correction. This study is under progress now.

The Poisson series is a new approach for fitting and describing the orbital prediction error, the benefits of the new approach include two aspects. Firstly, as error-fitting function, the Poisson series can describe variation of error with respect to propagation duration and on-orbit position of objects that can be described by true anomaly or mean anomaly. Due to the dynamics of the orbit and the orbit determination process, the orbital errors are strongly dependent on the on-orbit position (i.e. true anomaly or mean anomaly) of objects. HEO is the most but not the only affected orbit by this periodic characteristic. The dependence on on-orbit position of the prediction accuracy exists in all kinds of orbits, as Figures C.9–C.12 in Appendix C illustrated. However, most of current TLE error analysis methods use polynomial fitting which cannot represent the periodic characteristic of prediction error. The periodicity characterization and Poisson series fitting could improve accuracy of the orbit covariance information that can directly affect the validity of calculated collision probability. Secondly, in contrast to some approaches that consider periodic characteristics, the Poisson coefficient matrices can be obtained before close approach analysis, which is significant to timely assessment of collision risk. In addition, the orders of Poisson series are adjustable. The Poisson series is a common form for describing orbital prediction error.

The Poisson series fitting method is compatible with polynomial fitting methods developed by previous researchers. The commonly used polynomial fitting is a special case of the Poisson series fitting when the trigonometric coefficients vanish. The polynomial fitting is contained within the Poisson fitting results.

It is necessary to notice that the Poisson series fitting does not require any knowledge about how the state vectors are generated. It can thus be used for other orbit models and elements.

Table 3.3 The temporal and angular parameters of U.S.-Russian satellite collision

Satellite name	Cosmos-2251	Iridium-33
Epoch of TLE, t_0 (UTC)	2009-02-09 11:57:36.890	2009-02-09 18:49:39.280
Time of closest approach, t_{TCA} (UTC)	2009-02-10 16:55:59.796	2009-02-10 16:55:59.796
Prediction time, Δt (day)	1.20720956	0.92107078
Mean anomaly at TCA, M ($^\circ$)	13.167	243.232

Table 3.4 Position error standard deviations of two satellites

Satellite name	σ_U /km	σ_N /km	σ_W /km
Cosmos-2251	0.2061885	0.0231207	0.0719775
Iridium-33	0.4102069	0.0363234	0.0341134

3.6.3 Error Analysis on U.S.-Russian Satellite Collision

The first accidental collision of two intact spacecraft occurred on 10 February, 2009, leaving two distinct debris clouds extending through much of low Earth orbit (LEO) [31]. Iridium 33, a U.S. operational communications satellite (International Designator 1997-051C, U.S. Satellite Number 24946), and Cosmos 2251, a Russian decommissioned communications satellite (International Designator 1993-036A, U. S. Satellite Number 22675), collided at 16:55 UTC as the two vehicles passed over extreme northern Siberia at an altitude of 790 km. Both spacecraft were in nearly circular orbits with high inclinations: 86.4° and 74.04° , respectively.

The conjunction analysis gives the temporal parameters and angular parameters as shown in Table 3.3.

The TLE errors of two satellites are analyzed with the Poisson series fitting method of orbital prediction error. The propagation duration threshold $t_T = 2$ days. The maximum orders of Poisson series $n = 2$ and $m = 4$, so the Poisson coefficient matrices are (3×5) matrices. Substitution of the prediction time Δt and the mean anomaly M at TCA in Table 3.3 to the Poisson coefficient matrices yields the U, N, W direction position error standard deviations of two satellites at TCA as shown in Table 3.4.

Figures 3.9 and 3.10 illustrate the 3σ error ellipsoids of two satellites at TCA. The aspect ratios $\sigma_U : \sigma_N : \sigma_W$ of two ellipsoids are $9 : 1 : 3$ and $12 : 1 : 1$, respectively, which are close to the ratio $10 : 1 : 2$ in Ref. [32].

In this way the prediction error standard deviations of two satellites at TCA are gained. The collision probability could be calculated using the position and velocity vectors and the error information.

Fig. 3.9 3σ error ellipsoid of Cosmos-2251

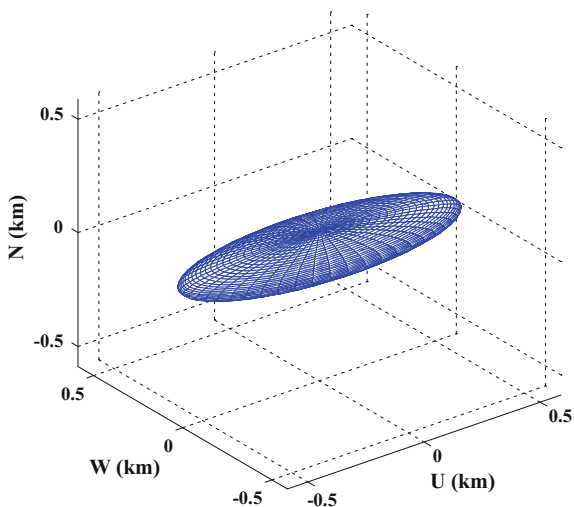
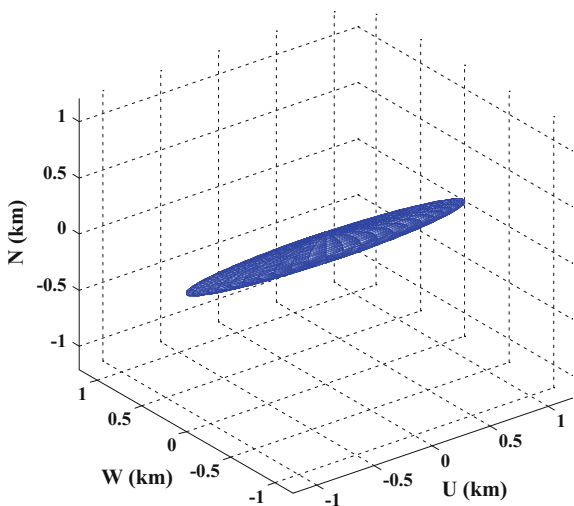


Fig. 3.10 3σ error ellipsoid of Iridium-33



3.7 Validation of Negative Correlation of Position and Velocity Error

This section will statistically analyze the TLE prediction residual data to validate the negative-correlation characteristics of position and velocity error presented in Chap. 2. The periodicity characteristics is not considered for simplicity because the negative-correlation characteristics of position and velocity error mainly involving

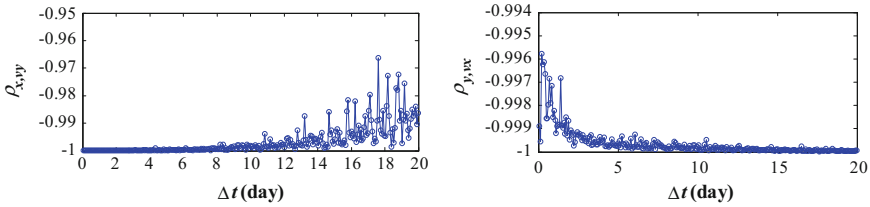


Fig. 3.11 Correlation coefficient curves of LEO object

the secular evolution. Only the orbit state at epoch of TLE chosen as the reference state when generate the residual data, i.e. the segment number $n_{seg} = 0$.

Let the temporal difference threshold $t_T = 20$ days, Figures C.13–C.15 in Appendix C illustrate the TLE’s prediction residual state components in UNW coordinate system versus temporal difference of three near-circular orbit LEO, MEO, and GEO.

Figures C.13–C.15 show that the position error in S direction and the velocity error in R direction in near-circular orbit have opposite propagation trends, so do the position error in R direction and the velocity error in S direction. The shapes of distributions of the position error in S direction and the velocity error in R direction are similar and symmetric about the temporal axis, so do the position error in R direction and the velocity error in S direction.

The correlation coefficient ρ is introduced in Sect. 2.4.2 to describe the opposite relation and similarity characteristics. Two correlation coefficient curves are provided by the propagation function of covariance based on the C-W Equation. This section validates the characteristic utilizing the TLE historical data. The residual data are 1-D binned according to the temporal parameters. The correlation coefficient $\rho_{x,vy}$ of R direction position error and S direction velocity error as well as the correlation coefficient $\rho_{y,vx}$ of S direction position error and R direction velocity error are calculated in each data bins. Figures 3.11, 3.12 and 3.13 illustrate the correlation coefficient curves versus prediction time.

Figures 3.11, 3.12 and 3.13 reveals that:

- (1) The correlation coefficient $\rho_{x,vy}$ of R direction position error and S direction velocity error ranges from -0.96 to -1 , a strong negative correlation exists

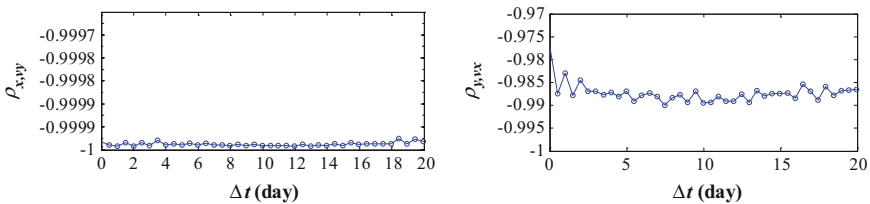


Fig. 3.12 Correlation coefficient curves of MEO object

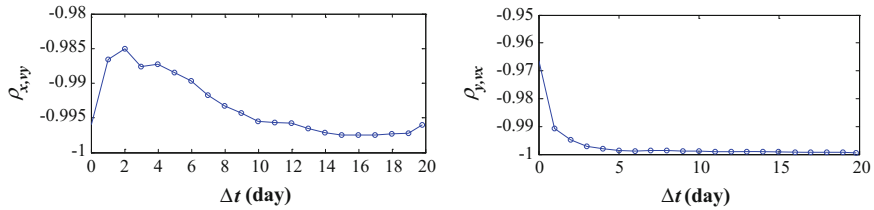


Fig. 3.13 Correlation coefficient curves of GEO object

between these two errors. However, the correlation coefficient has the trend to become positive for LEO object; this may be caused by the atmospheric drag.

- (2) The correlation coefficient $\rho_{y,vx}$ of S direction position error and R direction velocity error ranges from -0.97 to -1 , and has the trend to be -1 with the time increase. A strong negative correlation exists between these two errors.

References

1. Chan JC, Navarro D (2001) Comparisons of NORAD Two—line elements with INTELSAT orbital elements. In: Proceeding of the third european conference on space Debris. Darmstadt, Germany
2. Kelso TS (2007) Validation of SGP4 and IS—GPS-200D against GPS precision ephemerides [C]. AAS 07 -127. AAS/AIAA space flight mechanics meeting. Sedona, Arizona, 28 January-01 February 2007
3. Boyce WH (2004) Examination of NORAD TLE accuracy using the Iridium constellation. Spaceflight Mech. 119: 2133–2142 (Advances in the Astronautical Science)
4. Snow D, Kaya D (1999) Element set prediction accuracy assessment. Astrodynamics. 103: 1937–1958 (Advances in the Astronautical Science)
5. Krag H, Klinkrad H, Alarcon-Rodriguez JR (2007) Assessment of orbit uncertainties for collision risk predictions at ESA. In: Proceedings of second IAASS conference. Chicago, USA
6. Flohrer T, Krag H, Klinkrad H (2008) Assessment and categorization of TLE orbit errors for the US SSN catalogue. In: Proceedings of the advanced Maui optical and space surveillance technologies conference. Maui, US
7. Wang R, Liu J, Zhang QM (2009) Propagation errors analysis of TLE data. Adv Space Res 43:1065–1069
8. Peterson GE, Gist RG, Oltrogge DL (2001) Covariance generation for space objects using public data. Adv Astronaut Sci 1081:201–214
9. Deguine B, Foliard J, Alby F et al Covariance modeling in satellite collision risk activities. AIAA 2002-4631. In: AIAA/AAS astrodynamics specialist conference and exhibit. Monterey, California
10. Osweiler VP (2006) Covariance estimation and autocorrelation of NORAD two-line element sets. Air Force Institute of Technology, US Air University, Ohio
11. Legendre P, Deguine B, Garnier R et al (2006) Two line element accuracy assessment based on a mixture of Gaussian laws. AIAA 2006 - 6518. In: AIAA/AAS astrodynamics specialist conference and exhibit. Keystone, Colorado

12. Legendre P, Garmier R, Prat G et al (2007) Improvement of the two line element accuracy assessment based on a mixture of Gaussian laws. In: AAS 07 - 390. AAS/AIAA astrodynamics specialist conference. Mackinac Island, Michigan
13. Legendre P, Garmier R, Revelin B et al (2008) Improvement of the TLE accuracy model based on a Gaussian mixture depending on the propagation duration. In: AIAA 2008 -6772. AIAA guidance, navigation, and control conference and exhibit
14. Matney MJ, Anz-Meador P, Foster JL (2004) Covariance correlations in collision avoidance probability calculations. *Adv Space Res* 34:1109–1114
15. Hirose C, Kudo N, Matsuda I et al (2010) Evaluation of TLE prediction errors for conjunction assessment. IAC-10-A6.2.8. In: The 61st international astronomical congress. Prague, Czech
16. Levit C, Marshall W (2011) Improved orbit predictions using two-line elements. *Adv Space Res* 47:1107–1115
17. Muldoon AR, Elkaim GH, Rickard I F et al (2009) Improved orbital debris trajectory estimation based on sequential TLE processing. IAC -09. A6.2.9. In: The 60th international astronomical congress. Daejeon, R. Korea
18. Coppola VT, Woodburn J, Hujsak R (2004) Effects of cross correlated covariance on spacecraft collision probability. AAS 04-181. In: The AAS/AIAA space flight mechanics meeting
19. Hodge VJ, Austin J (2004) A survey of outlier detection methodologies. *Artif Intell Rev* 22:85–126
20. Jefferys WH (1970) A Fortran-based list processor for Poisson series. *Celest Mech* 2:474–480
21. Cherniack JR (1973) A more general system for Poisson series manipulation. *Celest Mech* 7:107–121
22. Ivanova T (2001) A new echeloned Poisson series processor (EPSP). *Celest Mech Dyn Astron* 80:167–176
23. Broucke R, Garthwaite K (1969) A programming system for analytical series expansions on a computer. *Celest Mech* 1:271–284
24. Henrard J (1989) A survey of Poisson series processors. *Celest Mech* 45:245–253
25. San-Juan JF, Abada A (2001) Algebraic and symbolic manipulation of Poisson series. *J Symbolic Comput* 32: 565–572
26. Martinez MC, Navarro JF, Ferrándiz JM (2007) An improved algorithm to compute circular functions of Poisson series. *Celest Mech Dyn Astron* 99:59–68
27. San-Juan JF, Abada A (2001) Communications between the Poisson series processor PSPC and general scientific software. *Math Comput Simul* 57: 307–315
28. Fateman RJ (1974) On the multiplication of Poisson series. *Celest Mech* 10:243–247
29. Fekken A, Hilhorst D, Kerckhoffs EJH et al (1996) Parallel computer algebra on Poisson series. *Simul Pract Theor* 4:219–243
30. Richardson DL (1989) PARSEC: an interactive Poisson series processor for personal computing systems. *Celest Mech* 45:267–274
31. Liou JC (2009) Satellite collision leaves significant debris clouds. *Orbital Debris Q News* 13 (2):1
32. Chan FK (2008) Spacecraft collision probability. The Aerospace Press, El Segundo

Chapter 4

Close Approach Analysis Between Space Object

The close approach analysis is to determine the time when two objects enter to the spherical or ellipsoidal range and the corresponding relative distance, the components of relative distance, the relative velocity and the conjunction angle, with the orbital data (orbital elements or high precision orbital ephemeris) of two objects are known and the close approach threshold (spherical or ellipsoidal threshold surface) and the time interval are predefined. The close approach analysis is not only significant in the conjunction assessment, but also used on different space mission analysis such as observing time determination of space-based space surveillance and reconstruction of spacecraft's maneuver.

The prefilter is needed to eliminate the objects that are impossible to conjunction the interesting spacecraft before further close approach analysis. The common prefilter methods include the apogee-perigee filter and altitude difference filter [1–3].

The close approach analysis methods can be categorized to the analytical method and the numerical method based on the orbital data and algorithm used. The analytical methods are based on the orbital elements and geometrical analysis of the orbital elements, obtain the close approach events' information by the analytical algorithm like derivative. The numerical methods are based on the orbital ephemeris of objects at certain time steps during a certain interval, or the position and velocity information from orbital model. The relative position and close approach information are obtained by numerical processing methods such as difference, interpolation, fitting, polynomial root-finding to the orbital ephemerides or position and velocity information.

The advantages of analytical methods are their computing speed and straightforward physical meanings, while the drawbacks are their sensitivity to the type of orbit and the type of orbital data and missing alarm risk due to the perturbations. The advantages of numerical methods are their relative low missing alarm risk and their robustness to the type of orbit, the type of orbital data, and the orbital models. The disadvantages are longer computing time. With the advent of ever increasing computer capacities numerical methods are more applicable in the future.

4.1 Analytical Methods to Determine Close Approach

4.1.1 Hoots's Method

Hoots proposed an analytical method to determine close approaches between satellites based on the geometrical filter in 1984 [3]. This method determines the close approaches by a series of altitude filter, orbital geometry filter and phase filter. This analytical filtering method is the early and mature algorithm, most of the other analytical methods are the improvement of Hoots's method [1, 4]. This detection concept for proximity events entails complex filter algorithms, with CPU-time demanding iterative root finders, and with a non-negligible risk of missing close fly-bys due to perturbations affecting the conjunction geometries. In addition, Refs. [5–10] studied the minimal distance and the critical points of the distance function between two confocal Keplerian orbit based on the orbit elements. This subsection will introduce the Hoots's method.

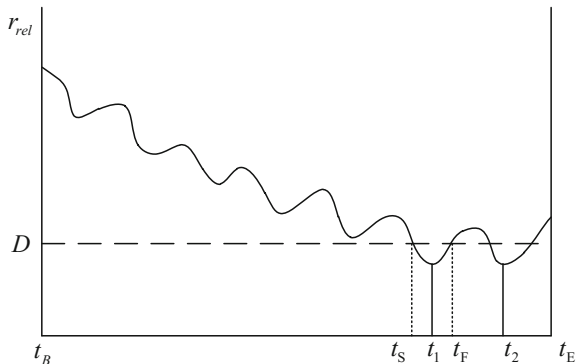
To protect or warn a satellite (called the primary) of a future collision, its orbit must be compared with every other object (called secondaries) in orbit about the Earth. The problem is stated most generally by specifying a time span of interest, $[t_B, t_E]$, and a separation distance D . As two satellites move around their respective orbits, the relative separation distance between the two is given by

$$r_{\text{rel}} = |\mathbf{r}_p - \mathbf{r}_s| \quad (4.1)$$

where \mathbf{r} denotes the position vector of the satellite and where the sub-scripts p and s denote the primary and secondary satellite, respectively.

Over the span of interest, a graph of relative distance might appear as in Fig. 4.1. We seek all points of closest approach which are less than the threshold D (e.g., times t_1 and t_2 in Fig. 4.1).

Fig. 4.1 Relative distance between two objects



Since these points are simply relative minima of the function $r_{\text{rel}}(t)$, the most obvious solution to the problem would be a numerical technique. Starting at the beginning of the time interval of interest, we could numerically step along the primary and secondary orbits noting the relative separation distances. Whenever the relative distance changed from decreasing to increasing, we would know we had passed a relative minima. If this minima was less than D , we would record this time as a time of close approach. Stepping would continue until we reached the end of the time interval of interest. If we wanted to find all times of close approach with all orbiting objects, the procedure would be repeated for each of the 15,000 secondaries. Clearly this procedure consumes exorbitant of computer processing time. When we consider the task of collision alerting for many primaries, use of numerical techniques becomes hopeless.

Thus, to make the task manageable, we take an analytical approach to the solution. The general problem is broken down into the non-coplanar and the coplanar cases. In each case, a list of candidate times is produced by analytical equations. These candidates are generally very close to the exact times of close approach. They are refined by an iterative technique to produce the times of closest approach. Periods of close proximity when the two satellites are within the distance D are calculated in a similar way.

In the discussion which follows, it will first be assumed that the satellites follow two-body Keplerian orbits with no perturbations. The effect of perturbations on the candidate times is small and is easily in-corporated as a minor modification to the analytical equations.

Geometrical Prefilter Many of the possible secondaries can be eliminated from consideration because the relative geometry of the two ellipses does not allow the two trajectories to ever come within the threshold distance D . The simplest test is a perigee-apogee test. Let q denote the larger of the two perigees and let Q denote the smaller of the two apogees. If

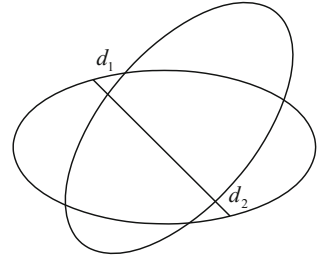
$$q - Q > D \tag{4.2}$$

then the secondary need not be considered further.

If a secondary is not eliminated by the perigee-apogee prefilter, then it is subjected to a second prefilter. This geometrical prefilter considers the relative geometry of the two ellipses in space. Figure 4.2 illustrates two arbitrary ellipses and the line of intersection of the two orbital planes.

There will be two values d_1 and d_2 of closest approach distance between the two elliptical paths. It should be noted that the points are not necessarily along the line of intersection of the two planes. This is a purely geometrical calculation and is independent of the location of the two satellites on their elliptical paths. If the minimum distance in space between the two paths is greater than the separation distance D , then the satellites cannot have a close encounter of interest. Let

Fig. 4.2 Geometrical prefilter



$$\mathbf{K} = \boldsymbol{\omega}_s \times \boldsymbol{\omega}_p \quad (4.3)$$

where $\boldsymbol{\omega}_s$ and $\boldsymbol{\omega}_p$ are unit vectors normal to the orbital planes of the secondary and primary satellites, respectively. These unit vectors can be calculated with the equation

$$\boldsymbol{\omega} = \sin \Omega \sin I \cdot \mathbf{i} + \cos \Omega \sin I \cdot \mathbf{j} + \cos I \cdot \mathbf{k} \quad (4.4)$$

where Ω is the longitude of ascending node, I is the inclination, and \mathbf{i}, \mathbf{j} , and \mathbf{k} are unit vectors in the X, Y, and Z directions.

The vector \mathbf{K} , lies along the line of intersection of the two orbital planes and has the additional property that

$$|\mathbf{K}| = \sin I_R \quad (4.5)$$

where I_R is the relative inclination angle between the two orbital planes. It should be noted that, in general, I_R is not simply the difference between the inclinations of the two orbits. If I_R is identically zero, the geometrical prefilter equations become indeterminate and such cases must be treated by the coplanar method.

The orientation of the vector \mathbf{K} within the orbital plane of either satellite is given by the angle Δ between the line of ascending node of the satellite plane and the vector \mathbf{K} as illustrated in Fig. 4.3. Let

$$\mathbf{n}_p = \cos \Omega_p \cdot \mathbf{i} + \sin \Omega_p \cdot \mathbf{j} \quad (4.6)$$

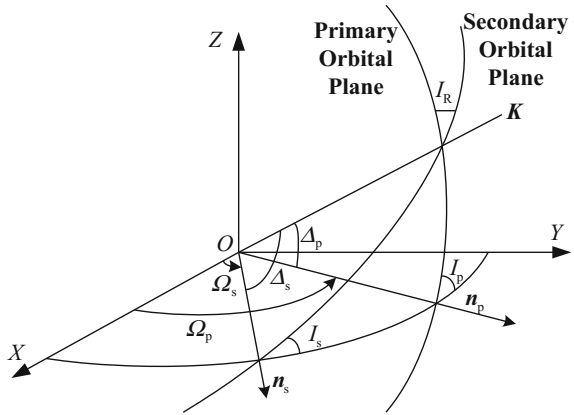
be a unit vector along the line of ascending node of the primary satellite. Then

$$\mathbf{K} \cdot \mathbf{n}_p = \sin I_p \cos I_s - \sin I_s \cos I_p \cos(\Omega_p - \Omega_s) \quad (4.7)$$

But from Fig. 4.3, it is clear that

$$\mathbf{K} \cdot \mathbf{n}_p = |\mathbf{K}| \cos \Delta_p = \sin I_R \cos \Delta_p \quad (4.8)$$

Fig. 4.3 Relative orientation of orbital planes



Thus,

$$\cos \Delta_p = \frac{1}{\sin I_R} [\sin I_p \cos I_s - \sin I_s \cos I_p \cos(\Omega_p - \Omega_s)] \quad (4.9)$$

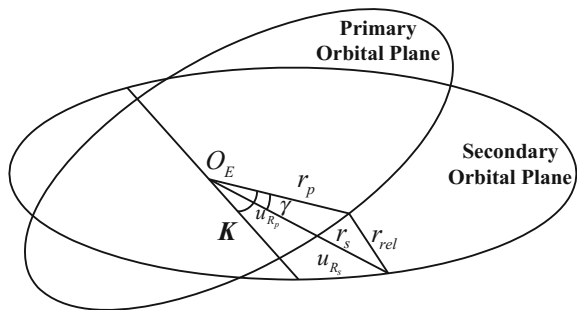
Also, by using the law of sines in spherical trigonometry, we get

$$\sin \Delta_p = \frac{1}{\sin I_R} \sin I_s \sin(\Omega_p - \Omega_s) \quad (4.10)$$

From Eqs. (4.9) and (4.10), Δ_p is uniquely determined. Similarly, it can be shown that

$$\begin{aligned} \cos \Delta_s &= \frac{1}{\sin I_R} [\sin I_p \cos I_s \cos(\Omega_p - \Omega_s) - \sin I_s \cos I_p] \\ \sin \Delta_s &= \frac{1}{\sin I_R} \sin I_p \sin(\Omega_p - \Omega_s) \end{aligned} \quad (4.11)$$

Fig. 4.4 Satellite positions in orbital planes



Let u_R denote the angle within the orbital plane from the vector \mathbf{K} to the satellite position measured positive in the sense of the satellite motion as shown in Fig. 4.4.

Let r_p and r_s denote the magnitude of the position vectors of the primary and secondary satellites and let γ denote the angle between the two position vectors. Then

$$r_{rel}^2 = r_p^2 + r_s^2 - 2r_p r_s \cos \gamma \quad (4.12)$$

From spherical trigonometry

$$\cos \gamma = \cos u_{R_p} \cos u_{R_s} + \sin u_{R_p} \sin u_{R_s} \cos I_R \quad (4.13)$$

The point of closest approach of the two elliptical paths will be a relative minimum of the function in Eq. (4.12). It is a function of the true anomalies of the two satellites since $u_R = f + \omega - \Delta$ where f is the true anomaly and ω is the argument of perigee. Thus, the relative minimum will be the solution of the two simultaneous equations

$$\frac{\partial r_{rel}^2}{\partial f_p} = 0, \quad \frac{\partial r_{rel}^2}{\partial f_s} = 0 \quad (4.14)$$

After considerable calculation and the use of Eqs. (4.13) and (4.14) reduce to

$$\begin{aligned} r_p e_p \sin f_p + r_s [\cos u_{R_s} (\sin u_{R_p} + a_{y_p}) - \sin u_{R_s} \cos I_R (\cos u_{R_p} + a_{x_p})] &= 0 \\ r_s e_s \sin f_s + r_p [\cos u_{R_p} (\sin u_{R_s} + a_{y_s}) - \sin u_{R_p} \cos I_R (\cos u_{R_s} + a_{x_s})] &= 0 \end{aligned} \quad (4.15)$$

where we have introduced the notation

$$\begin{cases} a_x = e \cos(\omega - \Delta) \\ a_y = e \sin(\omega - \Delta) \end{cases} \quad (4.16)$$

For the case of two circular orbits, Eq. (4.15) has the solution

$$\begin{aligned} u_{R_p} &= u_{R_s} = 0 \\ f_p &= \Delta_p - \omega_p, \quad f_s = \Delta_s - \omega_s \end{aligned} \quad (4.17)$$

For noncircular cases the analytical solution of Eq. (4.15) seems intractable. However, the equations readily admit an iterative solution by Newton's method. Given equations of the form

$$\begin{cases} F(f_p, f_s) = 0 \\ G(f_p, f_s) = 0 \end{cases} \quad (4.18)$$

Newton's method gives the iterative solution

$$f_{p_{i+1}} = f_{p_i} + h, \quad f_{s_{i+1}} = f_{s_i} + k \quad (4.19)$$

where

$$\left\{ \begin{array}{l} h = \frac{F \frac{\partial G}{\partial s} - G \frac{\partial F}{\partial s}}{\frac{\partial F \partial G}{\partial s \partial p} - \frac{\partial F \partial G}{\partial p \partial s}}, \quad k = \frac{G \frac{\partial F}{\partial p} - F \frac{\partial G}{\partial p}}{\frac{\partial F \partial G}{\partial s \partial p} - \frac{\partial F \partial G}{\partial p \partial s}} \\ F = r_p e_p \sin f_p + r_s (A \cos u_{R_s} - B \cos I_R \sin u_{R_s}) \\ G = r_s e_s \sin f_s + r_p (C \cos u_{R_p} - D \cos I_R \sin u_{R_p}) \\ \frac{\partial F}{\partial f_p} = r_p e_p \cos E_p + r_s \cos \gamma \\ \frac{\partial F}{\partial f_s} = -\frac{r_s}{1 + e_s \cos f_s} (AC + BD \cos I_R) \\ \frac{\partial G}{\partial f_p} = -\frac{r_p}{1 + e_p \cos f_p} (AC + BD \cos I_R) \\ \frac{\partial G}{\partial f_s} = r_s e_s \cos E_s + r_p \cos \gamma \end{array} \right. \quad (4.20)$$

with

$$\begin{array}{l} A = \sin u_{R_p} + a_{y_p}, \quad B = \cos u_{R_p} + a_{x_p} \\ C = \sin u_{R_s} + a_{y_s}, \quad B = \cos u_{R_s} + a_{x_s} \end{array} \quad (4.21)$$

where E denotes the eccentric anomaly and $\cos \gamma$ is given by Eq. (4.13). A starting value for the iteration is given by Eq. (4.17). Let f_{p^*} and f_{s^*} denote the final converged values from the iteration. These are used in Eq. (4.12) to calculate one point of closest approach between the two elliptical paths. There is another point of geometrical closeness which must be considered. It is found by iteration on Eq. (4.15) using the starting values.

$$\left\{ \begin{array}{l} f_{p1} = f_{p^*} + \pi \\ f_{s1} = f_{s^*} + \pi \end{array} \right. \quad (4.22)$$

The converged values from the iteration are used in Eq. (4.12) to calculate the other point of closest approach between the two elliptical paths. If both points of closest approach are greater than D , then the secondary need not be considered further.

Time Prefilter For those secondary satellites which have not been eliminated by the geometrical prefilters, there is yet one more constraint which can be used to remove them from consideration. Even though the elliptical paths may lie within distance D of each other, the two satellites must also simultaneously pass through these regions of closeness in order to have a close encounter between the satellites. This provides the means for a final prefilter based on crossing times for the two satellites of the line of intersection of the two orbital planes.

In general, a satellite will be in a region of vulnerability for a short period of time before and after it flies through the line of intersection of the two orbital planes. An estimate of these times can be obtained analytically by using the perpendicular distance z^* from one satellite to the other satellite's orbital plane. Referring to Fig. 4.4 we can see that will be less than or equal to r_{rel} ($z^* \leq r_{rel}$). Thus, the times of entry into and exit from the region of vulnerability may be slightly conservative. The perpendicular distance z^* can be determined with spherical trigonometry to be

$$z^* = r \sin I_R \sin u_R \quad (4.23)$$

Since we want to find the time when $z^* = D$, we make this substitution in Eq. (4.23) as well as a substitution for u_R to obtain

$$D = \frac{a(1 - e^2) \sin I_R \sin u_R}{1 + e \cos(u_R - \omega + \Delta)} \quad (4.24)$$

where a is the semimajor axis and e is the eccentricity. Using trigonometric identities, we can write

$$D = \frac{\alpha \sin u_R}{1 + a_x \cos u_R + a_y \sin u_R} \quad (4.25)$$

where $\alpha = a(1 - e^2) \sin I_R$ and a_x and a_y were defined in Eq. (4.16). Equation (4.25) can be solved for $\cos u_R$. After considerable algebra we find

$$\cos u_R = \frac{-D^2 a_x \pm (\alpha - D a_y) Q^{1/2}}{\alpha(\alpha - 2D a_y) + D^2 e^2} \quad (4.26)$$

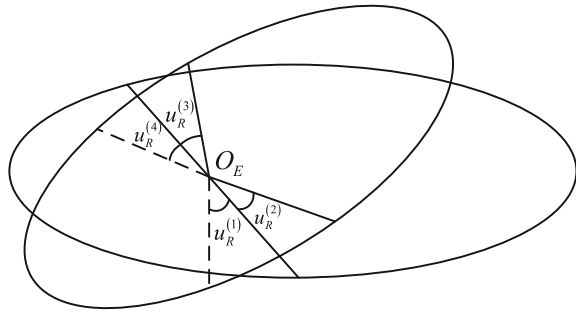
where

$$Q = \alpha(\alpha - 2D a_y) - (1 - e^2) D^2 \quad (4.27)$$

Now if Q is negative or the right-hand side of Eq. (4.26) has absolute value greater than one, it simply means that the satellite path does not ever get further than D from the other satellite's orbital plane. In such a case, the time prefilter cannot be applied and the satellite must be treated by the coplanar method.

If Eq. (4.26) is determinant for both the primary and secondary satellite orbital elements, we can generate angular windows for each satellite. By taking both values of the across in $[0, 2\pi]$ and by using both signs in Eq. (4.26), we obtain the two angular windows $[u_R^{(1)}, u_R^{(2)}]$ and $[u_R^{(3)}, u_R^{(4)}]$, as illustrated in Fig. 4.5.

Fig. 4.5 Angular windows



The angular windows can be converted to time windows by using Kepler’s equation to convert to mean anomaly and hence to time. The two time windows represent the times of vulnerability for the satellite over a single revolution. By adding multiples of the satellite period to the endpoints of each window, a sequence of time windows can be generated throughout the interval of interest $[t_B, t_E]$. This procedure is followed for both the primary and secondary satellites generating two sequences of time windows. The windows are then cross matched for possible overlapping. Any primary and secondary windows which overlap generate a candidate time calculated as the midpoint of the overlap. These candidate times are used as starting values for a iterative solution of the time of closest approach. This iteration will be discussed in a later section.

The time prefilter will generally eliminate many satellites from further consideration because their orbital motion is out of time phase. For those satellites that it does not eliminate, the interval of interest is usually shortened to a small subset of candidate times by the time prefilter.

Coplanar Satellites For those satellites which are exactly coplanar ($\sin I_R = 0$, $I_R = 0$ or $I_R = \pi$) or are so near coplanar that the time prefilter cannot be applied, we can generate candidate times for close approaches in the following manner. The relative separation distance is

$$r_{rel}^2 = r_p^2 + r_s^2 - 2\mathbf{r}_p \cdot \mathbf{r}_s \tag{4.28}$$

The time rate of change of this function is given by

$$R = r_p \dot{r}_p + r_s \dot{r}_s - \dot{\mathbf{r}}_p \cdot \mathbf{r}_s - \mathbf{r}_p \cdot \dot{\mathbf{r}}_s \tag{4.29}$$

Since R is proportional to the slope of the graph in Fig. 4.1 and since we are looking for relative minima of the graph, we merely start at the beginning of the interval of interest and time step along the primary and secondary orbits searching for a change in sign from negative to positive of the function R . If such a sign change occurs within a step, a candidate for a point of close approach is generated

based on linear interpolation. Since we are only searching for a sign change of the function, a large fraction of the orbital period can be used for the time step. In particular, 1/5 of the smaller of the two satellite's periods has proven to be quite adequate.

For those few satellites which fall into the coplanar category, this method provides an efficient way to generate time candidates for points of close approach.

Points of Close Approach and Periods of Close Proximity The candidate times generated either by the time window method or by the coplanar technique are usually quite close to the true times of minima. Hence, they provide very good starting values for a Newton's method solution for the zeros of the function R . The iteration is given by

$$t_{i+1} = t_i - \frac{R}{\dot{R}} \quad (4.30)$$

where

$$\dot{R} = \dot{r}_p^2 + r_p \ddot{r}_p + \dot{r}_s^2 + r_s \ddot{r}_s - \ddot{\mathbf{r}}_p \cdot \mathbf{r}_s - 2\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_s - \mathbf{r}_p \cdot \ddot{\mathbf{r}}_s \quad (4.31)$$

Newton's method for this function is very reliable and converges quite rapidly to give the time of closest approach.

When two satellites have a close encounter less than D distance, it is of interest to know the beginning and end of the time interval for which they were within D distance of each other (e.g., times t_S and t_F in Fig. 4.1). That is, we seek two zeros of the function

$$H = r_{rel}^2 - D^2 \quad (4.32)$$

which are in the neighborhood of the minimum point. Since $\dot{H} = 2R$ from Eqs. (4.28) and (4.29). Newton's method gives

$$t_{i+1} = t_i - \frac{H}{2R} \quad (4.33)$$

A starting value for the beginning time of close proximity is taken as a few minutes prior to the minimum time point, and a starting value for the end time of close proximity is taken as a few minutes after the minimum time point.

4.1.2 Collision Detection Based on Secular Orbital Model

Owing to perturbations the motion of a space object is not a simple periodical one. An analytical method of collision detection in which the secular effects of

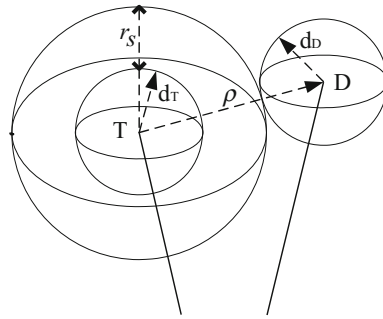


Fig. 4.6 The value of the alert range ρ

Table 4.1 Orbital types and parameters of objects

Number	Type	Inclination/(°)	Eccentricity	Perigee/km	Apogee/km
28883	LEO	42.471	0.002161	294.720	324.018
23560	SSO	98.5690	0.0000917	781.777	799.058
25030	MEO	56.277	0.010179	19914.860	20452.619
29260	HEO	63.7889	0.6804348	2120.177	38276.472
26038	GEO	0.068	0.000036	35786.059	35790.681
07578	GSO	14.4139	0.0002164	35865.438	35880.355

atmospheric drag and gravitation are included is discussed in this section. This algorithm is distinguished from Hoots’s method at the consideration of the secular effects of perturbations, which is useful for collision avoidance maneuver.

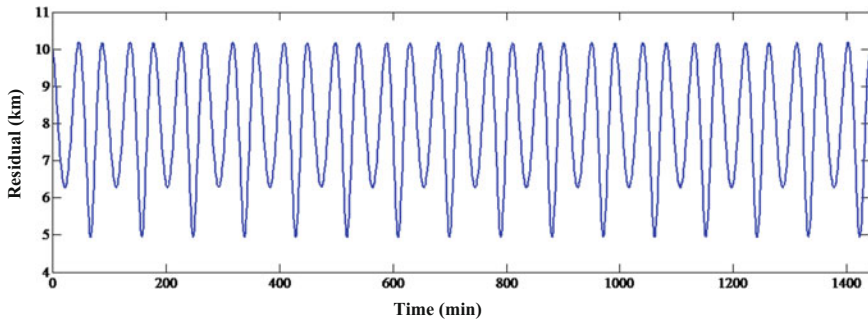
Precision of SGP4/SDP4 Model Only Including Secular Effects The orbital prediction is the basis of the collision detection method, and it also determines the alert range ρ , as shown in Fig. 4.6.

Supposing a spacecraft is T and a piece of debris is D, d_T is the orbital prediction precision of the spacecraft, d_D is the one of the debris and r_s is the safe distance relative to the special size and mission of the spacecraft. The alert range is

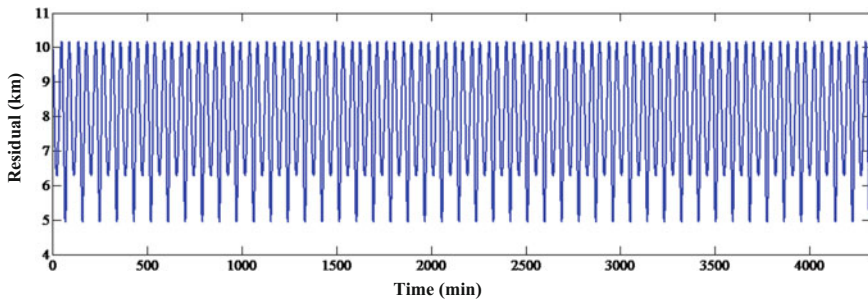
$$\rho = d_T + r_s + d_D \tag{4.34}$$

Equation (4.34) indicates that the prediction precision of orbital model directly affects the alert range.

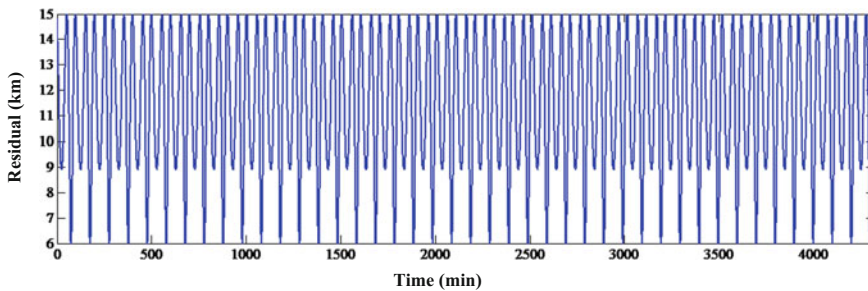
At present the orbital data of the space debris mainly come from the Two-Line Element (TLE) sets periodically issued by JSpOC. The TLE uses orbital propagator SGP4/SDP4 to obtain a position and velocity of the space object. The SGP4/SDP4 propagator considered the secular, the long-period and the short-period effects of gravitation and atmospheric drag etc. The secular effects could be computed in the polynomial functions, while the long-period and the short-period effects would be



(a) LEO object 28883 (1 day)



(b) LEO object 28883 (3 day)

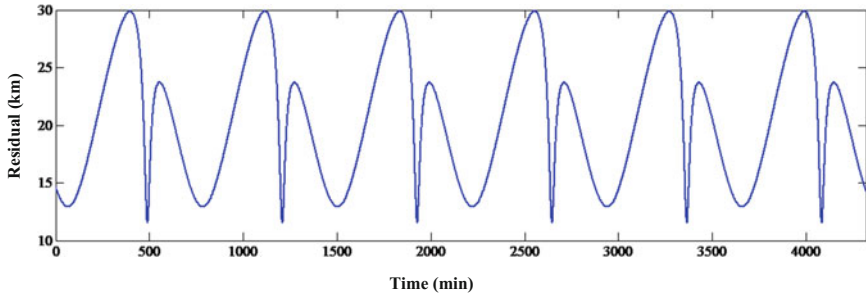


(c) SSO object 23560 (3 day)

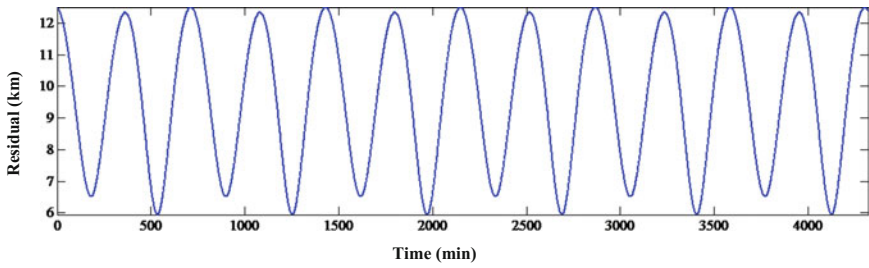
Fig. 4.7 Residual of the SGP4 model including only the secular terms

computed by iteration. In order to realize the quick collision detection the long-period and the short-period effects should be omitted during the process. The effects of the mitting of long-period and the short-period terms to the prediction precision are simulated in this section.

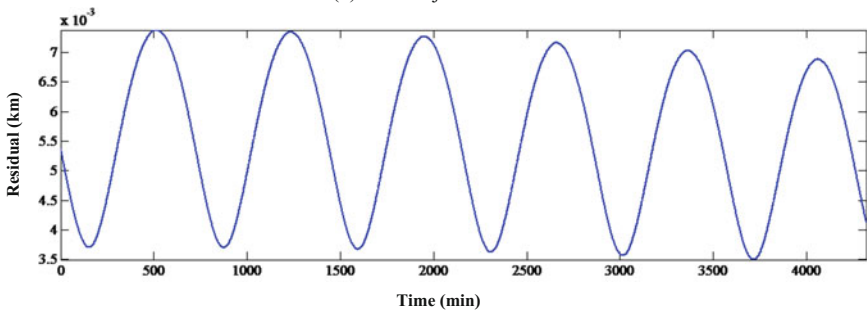
Six typical orbiting objects are selected for examples, the orbital types and parameters are shown in Table 4.1. The residuals of predicted position using the SGP4/SDP4 without the periodic terms are illustrated in Figs. 4.7 and 4.8.



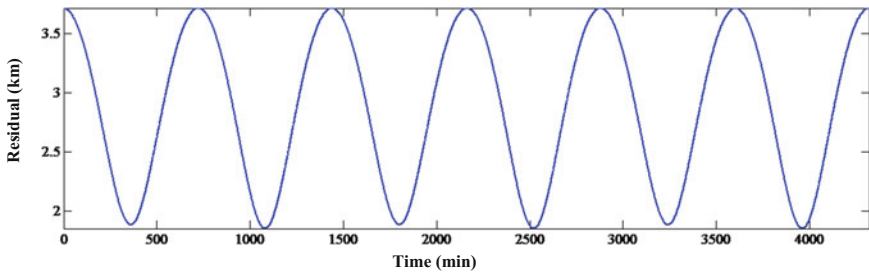
(a) HEO object 29260



(b) MEO object 25030



(c) GEO object 26038



(d) GSO object 07578

Fig. 4.8 Residual of the SDP4 model including only the secular terms

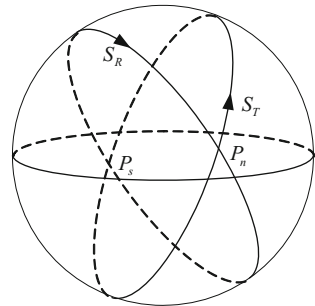
Figures 4.7 and 4.8 reveal that:

- (1) The oscillation of the position residual is mainly due to the short-period terms. The short-period terms are of the magnitude of orbital period or less, so the periods of the oscillation of LEO's residual are less than the ones of deep space orbits.
- (2) The atmospheric drag mainly affects the long-term evolution of orbit for the LEO object 28883, the oscillation limited in 10 km in 3 day prediction period. The inclination of the SSO object 23560 is close to 90° , the periodic terms of the SGP4 model include $\sin i$ and $(1 - \cos^2 i)$ which approximately equal to one when $i \approx 90^\circ$, therefore the effect is more significant than object 28883.
- (3) The eccentricity and inclination of the HEO object 29260 are $e = 0.68$ and $i = 63.8^\circ$, respectively. The effects of e and i related terms in SGP4 model are enhanced. The residual near the perigee changes remarkably.
- (4) The inclination and eccentricity of GEO are close to zero, the e and i related terms are insignificant, and the residual of position are of the magnitude of meters. The inclination of the GSO has drifted to 14° , which causes more periodic effects than GEO orbit, and the residual are of the magnitude of kilometers.
- (5) The residuals of position of the SGP4/SDP4 model including only the secular effects of the perturbation for 3 days prediction are limited in 30 km, which indicates it is reasonable to only take the secular term into consideration for the primary filter of close approach analysis.

Analytic Algorithm of Collision Detection Based on Secular Effects The algorithm is used to estimate the possibility of collision among spacecrafts and debris. Estimate rules are determined by the predicting precision of secular effects. The algorithm consists of two steps. First, the intersection points of the spherical projection of two orbits are analyzed. Then if the times crossing the points are close to each other, the altitudes of two objects are investigated to determine if these two objects tend to be close approach.

There are two points of intersection obtained from spherical projection of two orbits. One point is north of the equator and called northern point P_N , the other is south of the equator and called southern point P_S . The spherical angle between

Fig. 4.9 Points of intersection of spherical projection of two orbits



these two points is close to 180° (perturbation results in the angle difference to 180°). These points appear periodically for each orbit, as illustrated in Fig. 4.9. Collision could be detected effectively as long as the relationship between two objects at these points was obtained. Northern point and southern point are moving because of perturbation. As a result, northern point at last period probably turns to southern point at this period, and southern point at last period turns to northern point. This results in angle modification and must be carefully treated.

Supposing there are two objects in orbit, the first object's mean elements at time t_{01} are

$$\mathbf{X}_{01} = [a_{01} \quad e_{01} \quad i_{01} \quad \Omega_{01} \quad \omega_{01} \quad M_{01}]^T \tag{4.35}$$

According to the formulas listed in Refs. [3, 11], the orbital elements at any time t can be obtained with the help of secular effects.

$$\begin{aligned} a_1(t) &= a_{01} + g_{a1}(t - t_{01}) & e_1(t) &= e_{01} + g_{e1}(t - t_{01}) \\ i_1(t) &= i_{01} + g_{i1}(t - t_{01}) & \Omega_1(t) &= \Omega_{01} + g_{\Omega1}(t - t_{01}) \\ \omega_1(t) &= \omega_{01} + g_{\omega1}(t - t_{01}) & M_1(t) &= M_{01} + g_{M1}(t - t_{01}) \end{aligned} \tag{4.36}$$

where $g_{a1}(t - t_{01})$, $g_{e1}(t - t_{01})$, $g_{i1}(t - t_{01})$, $g_{\Omega1}(t - t_{01})$, $g_{\omega1}(t - t_{01})$, $g_{M1}(t - t_{01})$ are functions of time t , represents the effects of the perturbation to the elements.

Similarly, the secondary object's mean elements at time t_{02} are

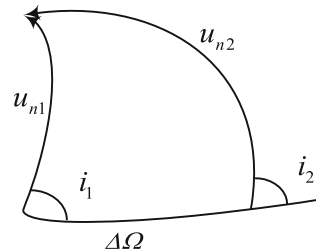
$$\mathbf{X}_{02} = [a_{02} \quad e_{02} \quad i_{02} \quad \Omega_{02} \quad \omega_{02} \quad M_{02}]^T \tag{4.37}$$

the orbital elements at any time t are

$$\begin{aligned} a_2(t) &= a_{02} + g_{a2}(t - t_{02}) & e_2(t) &= e_{02} + g_{e2}(t - t_{02}) \\ i_2(t) &= i_{02} + g_{i2}(t - t_{02}) & \Omega_2(t) &= \Omega_{02} + g_{\Omega2}(t - t_{02}) \\ \omega_2(t) &= \omega_{02} + g_{\omega2}(t - t_{02}) & M_2(t) &= M_{02} + g_{M2}(t - t_{02}) \end{aligned} \tag{4.38}$$

Northern point P_N is shown in Fig. 4.10, u_{n1} and u_{n2} can be obtained as below

Fig. 4.10 Northern point of intersection of spherical projection of two orbits



$$\begin{aligned}
u_{n1} &= \operatorname{arccot} \frac{\cos \Delta\Omega \cos i_{n1} - \sin i_{n1} \cot i_{n2}}{\sin \Delta\Omega} \\
u_{n2} &= \operatorname{arccot} \frac{-\cos \Delta\Omega \cos i_{n2} + \sin i_{n2} \cot i_{n1}}{\sin \Delta\Omega} \\
\Delta\Omega &= \Omega_{n2} - \Omega_{n1}
\end{aligned} \tag{4.39}$$

where Ω_{n1} , i_{n1} , Ω_{n2} , i_{n2} are orbital element of two objects when passing the northern point.

$$\begin{cases} f_{n1} = u_{n1} - \omega_{n1} \\ \tan \frac{E_{n1}}{2} = \sqrt{\frac{1-e_{n1}}{1+e_{n1}}} \cdot \tan \frac{f_{n1}}{2} \\ M_{n1} = E_{n1} - e_{n1} \cdot \sin E_{n1} \end{cases} \tag{4.40}$$

Let

$$\kappa_{n1}(\mathbf{X}_{01}, \mathbf{X}_{02}, t) = M_{n1} - M_1 = M_{n1} - M_{01} - g_{M1}(t - t_{01}) \tag{4.41}$$

When $\kappa_{n1}(\mathbf{X}_{01}, \mathbf{X}_{02}, t) = 0$, the corresponding t_{N1} is the time S_1 passing the northern point P_N . Similarly, let

$$\kappa_{n2}(\mathbf{X}_{01}, \mathbf{X}_{02}, t) = M_{n2} - M_2 = M_{n2} - M_{02} - g_{M2}(t - t_{01}) \tag{4.42}$$

When $\kappa_{n2}(\mathbf{X}_{01}, \mathbf{X}_{02}, t) = 0$, the corresponding t_{N2} is the time S_2 passing the northern point P_N .

Let $\Delta t_N = t_{N2} - t_{N1}$ be the time difference of S_1 and S_2 crossing the northern point. Notice that Δt_N is not zero in general. S_1 and S_2 pass the northern point P_N simultaneously when $\Delta t_N = 0$. If the radial of two orbits are same, S_1 and S_2 will collide at the northern point P_N .

The formulae at the southern point P_S are similar to the ones at P_N .

$$\begin{aligned}
u_{s1} &= \operatorname{arccot} \left(\frac{\cos \Delta\Omega \cos i_{s1} - \sin i_{s1} \cot i_{s2}}{\sin \Delta\Omega} \right) + \pi \\
u_{s2} &= \operatorname{arccot} \left(\frac{-\cos \Delta\Omega \cos i_{s2} + \sin i_{s2} \cot i_{s1}}{\sin \Delta\Omega} \right) + \pi \\
\Delta\Omega &= \Omega_{s2} - \Omega_{s1}
\end{aligned} \tag{4.43}$$

Besides, these special conditions below must be considered individually:

- (1) $i_1 > i_2 > 0$, $\Delta\Omega = \Omega_2 - \Omega_1 = 0$ or $\pm\pi$, the intersection points are their ascending node or descending node, the times of two objects passing the ascending or descending node are needed.
- (2) $i_1 > i_2 = 0$, the times of S_1 passing the ascending or descending node are needed.

Table 4.2 TLEs of two space objects

Object	TLE
S_R	1 26207U 99057CV 05012.01348642 .00000774 00000-0 27275-3 0 606
	2 26207 98.2173 32.2878 0124063 79.0395 282.4771 14.33127993251122
S_T	1 07219U 74015B 05005.94102452 .00000041 00000-0 37371-4 0 3267
	2 07219 99.0916 339.3105 0066266 134.9684 225.6893 14.24160453598108

(3) $i_1 = i_2$, $\Delta\Omega = \Omega_2 - \Omega_1 = 0$ or $\pm\pi$, two orbits are coplanar, the coplanar method to close approach analysis is needed.

Collision may happen if these two objects pass the point of intersection at the same time. Thus collision can be detected by comparing time t_1 and t_2 computed above. Collision could not happen if $|t_1 - t_2| > \varepsilon_t$, otherwise

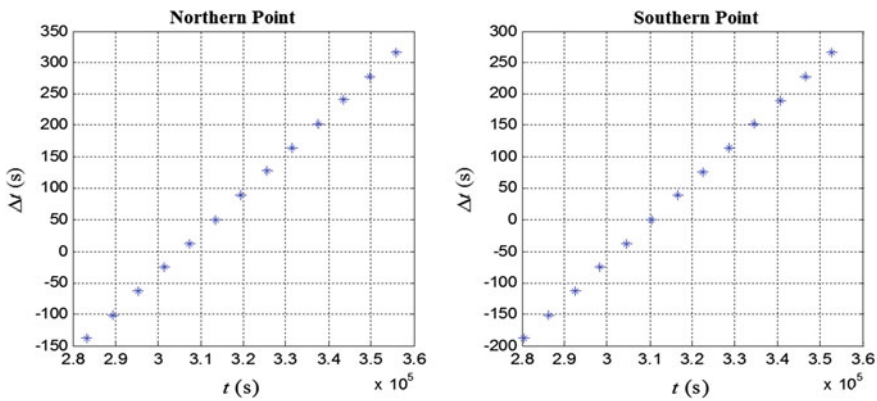
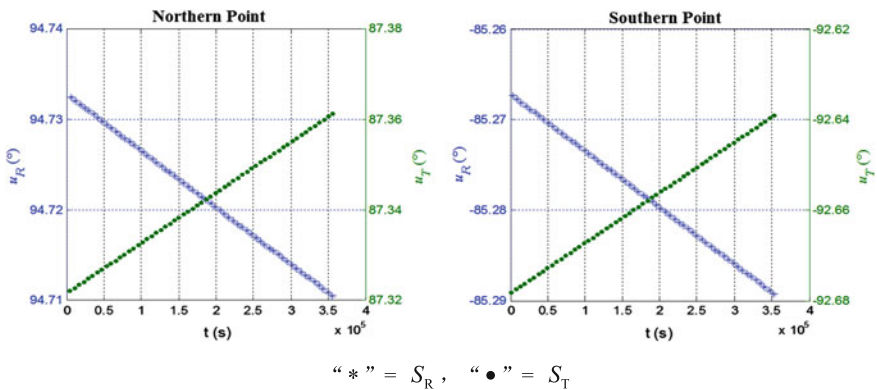


Fig. 4.11 The time differences of two objects crossing the northern point and southern point



“*” = S_R , “•” = S_T

Fig. 4.12 The orbital elements two objects crossing the northern point and southern point

$$\begin{cases} r_1 = a_1(1 + e_1 \cos E_1) \\ r_2 = a_2(1 + e_2 \cos E_2) \end{cases} \quad (4.44)$$

Collision could not happen if $|r_1 - r_2| > \varepsilon_r$ otherwise collision warning information must be issued. ε_t and ε_r are determined by precision of secular effects model.

The collision event in Ref. [12] is taken as an example. The two-line elements of these objects (shown in Table 4.2) are provided by NORAD. Supposing SSN 26207 is S_R , SSN 07219 is S_T . The trends of time crossing the northern and southern points and orbital elements at P_N and P_S are listed below. The initial epoch is 2005-01-13 12:00:00, the duration length is 4 days.

The time differences of two objects crossing P_N and P_S are shown in Fig. 4.11. The time differences Δt at P_N and P_S change discontinuously. Collision could be avoided if the time differences went away from zero with the help of maneuver. There is one time difference equal to zero in Fig. 4.11, which indicates there will be a conjunction. Fig. 4.12 illustrates the trend of the points of intersection of two orbits in 4 days.

4.2 Numerical Methods to Determine Close Approach

With the advent of ever increasing computer capacities numerical algorithms for conjunction event detections have become available. The numerical methods to determine close approach are based on the orbital ephemeris of objects at certain time steps during a certain interval, or the position and velocity information from orbital model. The relative position and close approach information are obtained by numerical processing methods such as difference, interpolation, fitting, polynomial root-finding to the orbital ephemerides or position and velocity information.

An obvious solution to the time and distance portion of the close approach problem is to step sequentially along the orbits of the two objects to form an ephemeris for each, and then difference the position vectors to determine their relative distance. Indeed, this approach is used to construct a truth table of encounters for analyses. Although straightforward, this approach requires many computations, especially for a lot of objects.

Alfano and Negron developed a numerical methods based on polynomial splining techniques to determine close approach (A-N algorithm) [13, 14]. The principle of A-N algorithm is to introduce distance function and ellipsoidal function, the local minimum relative distances and the corresponding times, as well as the time when one space object enters and exits the ellipsoidal region of another one are determined by solving the roots of the polynomial. This method determines close approaches between two arbitrary trajectories, and is not limited to a specific analytical, numerical, or hybrid propagator. It has several advantages over

traditional geometrical and brute-force method. A-N algorithm uses a relative-distance function, for which each point and slope matches in the blending functions. The method uses any propagation scheme, permitting us to select a desired speed and accuracy. The answers result from closed-form solutions and eliminate the need for iteration.

However, in A-N algorithm there are limitations in the criteria used to check root existence and the process of filtrating rational real roots. Furthermore, the criteria proposed in A-N algorithm for determination of the time step of splining are not appropriate for direct calculation and may lead to root loss. Li Jian corrected these insufficiencies [15]. A self-adaptive method based on polynomial splining error theory for choosing interpolation step is presented. Comparison shows that the result of the improved A-N algorithm is more reliable than the original one.

Alarcon-Rodriguez has devised a “smart sieve” conjunction detection algorithm based on fundamental flight dynamics principles [16, 17]. This method presents neither constraints nor singularities for any terrestrial orbit. This “smart sieve” algorithm consists of an initial altitude range filter, a crude range filter (based on the maximum possible velocities) for each of the three coordinates of the relative position vector, and three steps of refined distance filter (adjusted for actual range, maximum relative acceleration, and actual relative velocity), and a final range-rate root finder to determine the time of closest approach. Due to their simplicity and robustness this sieve algorithm is superior in their CPU time efficiency, and in their conjunction event detection capability.

Wang Ting combined the A-R dynamics sieve algorithm and A-N algorithm [18], the computing efficiency and stability increase evidently. Faulds and Spencer developed an algorithm with a low-order propagator, incorporating parallel processing techniques and a genetic algorithm to find closest approach [19].

4.2.1 A-R Algorithm

As part of a collision risk assessment one first needs to determine events of near misses between pairs of space objects from the trackable catalog population. The orbits of the operational satellites will be assumed to be available from an operational orbit determination at epoch t_0 , with resulting orbit files containing fitted and predicted states for a certain timespan. The operational satellites will hereafter be denoted as targets (index “t”). The potential conjunction objects are extracted from the USSTRATCOM catalog, with recent orbit information provided in TLE format. These TLE data provide first-order osculating states by means of the SGP4/SDP4 orbit theory. The conjunction counterparts will be denoted as risk objects (index “r”).

Different methods have been devised to determine close conjunction events between pairs of objects during a pre-defined time interval. Hoots’s conjunction detection is performed by the successive application of an altitude filter, a plane

geometry filter, and a phase filter. This approach is particularly useful, if the two orbits are available through an analytic orbit theory. This detection concept for proximity events entails complex filter algorithms, with CPU-time demanding iterative root finders, and with a non-negligible risk of missing close fly-bys due to perturbations affecting the conjunction geometries. With ever-increasing computer capacities alternative algorithms for conjunction event detections became a viable alternative. The following sieve algorithm has been implemented in ESA's CRASS software [16, 17].

As first step of the sieve, again an altitude filter is applied to the two orbit histories of the target and risk object. In a subsequent step, the time history of ranges $\rho(t)$ between the target orbit $\mathbf{r}_t(t)$ and the risk orbit $\mathbf{r}_r(t)$ is analyzed at equidistant time steps Δt across the prediction interval (for instance $t \in [t_0, t_0 + 7\text{d}]$). At each time step the coordinates of the range vector are checked against an adjusted safety distance

$$R_{c,1} = R_c + v_e \Delta t \quad (4.45)$$

where $v_e = \sqrt{2\mu/r}$ is the escape velocity. A close conjunction within a time interval Δt is not possible, if $\rho_X > R_{c,1}$, or $\rho_Y > R_{c,1}$, or $\rho_Z > R_{c,1}$, or $\rho > R_{c,1}$ (checked in this sequence). For those orbits which passed this step, the reference control volume of radius R_c around the target object cannot be violated, if a curvature adjusted safety distance $R_{c,2} = R_c + g\Delta t^2$ is not penetrated (where g is the gravity acceleration of the Earth). This corresponds to the following condition.

$$R_{c,2} = R_c + g\Delta t^2 < \sqrt{\Delta r^2 - (\Delta \mathbf{r} \cdot \Delta \mathbf{v} / \Delta v)^2} \quad (4.46)$$

The threshold $R_{c,1}$ can be further refined, considering the effective relative velocity, rather than the escape velocity, to obtain

$$R_{c,3} = R_{c,2} + \frac{1}{2} |\Delta \mathbf{v} \cdot \Delta \mathbf{r} / \Delta r| \Delta t \quad (4.47)$$

For those orbits passing the previous range sieves a root finder is started, which determines the time t_{tca} of closest approach $\Delta r_{tca} = \rho_{tca}$ from a zero-transition of the range-rate time history.

$$\Delta \mathbf{v}_{tca} \cdot \Delta \mathbf{r}_{tca} / \Delta r_{tca} = \dot{\rho}_{tca} = 0.0 \quad \rightarrow \quad t_{tca} \quad (4.48)$$

For each of the identified conjunctions a final test is performed to verify that $\Delta r_{tca} \leq R_c$. At the established time t_{tca} and for the associated target state $\mathbf{x}_t = (\mathbf{r}_t, \mathbf{v}_t)$ a target-centered, orbit-related coordinate system \mathbf{U}_t , \mathbf{V}_t and \mathbf{W}_t can be deployed according. This leads to a transformation matrix $\mathbf{R}_{t,UVW}$, which allows a mapping (X, Y, Z) to (U, V, W) of the conjunction distance $\Delta \mathbf{r}_{tca}$, and approach velocity $\Delta \mathbf{v}_{tca}$ in the case of a direct hit.

$$\begin{aligned}
(\Delta \mathbf{r}_{tca})_{U,V,W} &= (\Delta \mathbf{r}_t)_{U,V,W} = (\mathbf{R}_t)_{U,V,W} (\Delta \mathbf{r}_{tca})_{X,Y,Z} \\
(\Delta \mathbf{v}_{tca})_{U,V,W} &= (\Delta \mathbf{v}_t)_{U,V,W} = (\mathbf{R}_t)_{U,V,W} (\Delta \mathbf{v}_{tca})_{X,Y,Z}
\end{aligned} \tag{4.49}$$

where $\Delta \mathbf{r}_{tca}$ and $\Delta \mathbf{v}_{tca}$ are defined as

$$\begin{aligned}
(\Delta \mathbf{r}_{tca})_{X,Y,Z} &= (\mathbf{r}_r(t_{tca}) - \mathbf{r}_t(t_{tca}))_{X,Y,Z} \\
(\Delta \mathbf{v}_{tca})_{X,Y,Z} &= (\mathbf{v}_r(t_{tca}) - \mathbf{v}_t(t_{tca}))_{X,Y,Z}
\end{aligned} \tag{4.50}$$

The vector $(\Delta \mathbf{r}_{tca})_{U,V,W}$ can be expressed in terms of distance Δr_{tca} , azimuth $A_{r,tca}$, and elevation $h_{r,tca}$ of the conjunction. Correspondingly, the approach velocity $(\Delta \mathbf{v}_{tca})_{U,V,W}$ of the risk object can be expressed in terms of its magnitude Δv_{tca} , and the azimuth $A_{v,tca}^* = A_{v,tca} + \pi$, and elevation $h_{v,tca}^* = -h_{v,tca}$ of the direction from which the risk object would approach in the case of a direct hit on the target.

So far, a spherical threshold surface has been deployed around the target. Due to the secular effects of drag perturbations on the orbit period, with a resulting dominance of along-track position uncertainties, it is advantageous to define an ellipsoidal threshold surface, with its major axis in the along-track direction. The CRASS program uses an ellipsoid aligned with the \mathbf{U}_t , \mathbf{V}_t and \mathbf{W}_t axes, with dimensions $R_{c,U} = 10$ km, $R_{c,V} = 25$ km, and $R_{c,W} = 10$ km, for a spherical threshold radius of $R_c \geq \max(R_{c,U}, R_{c,V}, R_{c,W})$ used in the sieve algorithms. A conjunction event falls into the ellipsoidal control volume, if $k_c^2 \leq 1$ in the following equation. Else, the event is rejected from further analysis.

$$k_c^2 = \left(\frac{\Delta r_U}{R_{c,U}} \right)^2 + \left(\frac{\Delta r_V}{R_{c,V}} \right)^2 + \left(\frac{\Delta r_W}{R_{c,W}} \right)^2 \tag{4.51}$$

Owing to more frequent calls to the sieve algorithms, the computer run-time for the conjunction event sieve increases for steps that are too small. The run-time also increases for time steps that are too large, since fewer orbits can be rejected by the CPU-time-efficient initial sieve steps. In between these extremes one can find an optimal analysis time step, which is on the order of $\Delta t \approx 180$ s (about 3% of the orbital period of ERS-2 and Envisat).

Due to their simplicity and robustness the outlined sieve algorithms are found to be superior in their CPU time efficiency, and in their conjunction event detection capability, as compared with traditional methods which are based on filters considering altitude, orbit plane intersection, orbit phasing, and minimum approach distance.

4.2.2 A-N Algorithm

The method developed by Alfano and Negron [13, 14] is called the A-N algorithm. This method determines close approaches between two arbitrary trajectories. It has several advantages over traditional geometrical and brute-force method. A-N algorithm uses a relative-distance function, for which each point and slope matches in the blending functions. The main goal is determining the distance between the two satellites. The method uses any propagation scheme, permitting us to select a desired speed and accuracy. Finally, the answers result from closed-form solutions and eliminate the need for iteration.

The first calculation compares the apogee and perigee values of the orbits. This provides a rough cut to determine if close approach are possible. If the absolute value between the largest perigee and the smallest apogee is greater than the desired distance tolerance, the two objects will never approach one another, unless a thrusting maneuver is applied.

The position vectors of the primary object and secondary object in the geocentric equatorial (ECI) are \mathbf{r}_p and \mathbf{r}_s at time t . Here the inertial vectors are used, although Earth-fixed vectors are feasible too. The relative-distance vector \mathbf{r}_d , and its time derivatives become

$$\begin{cases} \mathbf{r}_d = \mathbf{r}_s - \mathbf{r}_p \\ \dot{\mathbf{r}}_d = \dot{\mathbf{r}}_s - \dot{\mathbf{r}}_p \\ \ddot{\mathbf{r}}_d = \ddot{\mathbf{r}}_s - \ddot{\mathbf{r}}_p \end{cases} \quad (4.52)$$

where any orbit propagator can produce the ephemerides for the primary and secondary objects. Large time steps with the propagator can be taken because of the splining techniques. If the objects' acceleration vectors aren't available, the vector can be approximated using the following equation includes only J_2 perturbative force.

$$\mathbf{a} = -\frac{\mu}{r^3}\mathbf{r} + \mathbf{a}_{\text{nonspherical}} + \mathbf{a}_{\text{drag}} + \mathbf{a}_{3\text{-body}} + \mathbf{a}_{\text{SR}} + \mathbf{a}_{\text{tides}} + \mathbf{a}_{\text{other}} \quad (4.53)$$

We define the distance function $f_d(t)$ and its time derivatives using dot products. We define the distance function to be the square of the distance so we don't need to evaluate a square root.

$$\begin{cases} f_d(t) = \mathbf{r}_d \cdot \mathbf{r}_d \\ \dot{f}_d(t) = 2(\dot{\mathbf{r}}_d \cdot \mathbf{r}_d) \\ \ddot{f}_d(t) = 2(\ddot{\mathbf{r}}_d \cdot \mathbf{r}_d + \dot{\mathbf{r}}_d \cdot \dot{\mathbf{r}}_d) \end{cases} \quad (4.54)$$

We evaluate these equations for a sequence of times until two adjoining times (containing a minimum) are found. Object close approaches occur whenever $f_d(t)$ is at a local minimum, when $\dot{f}_d(t) = 0$ and $\ddot{f}_d(t) > 0$. The final condition implies that

the two objects are in the same half plane of the orbit. To determine these times of closest approach, compute the coefficients α_{cd} of the derivative function for the range-rate cubic polynomial equation $P_{cd}(\tau)$ that corresponds to $\dot{f}_d(t)$. Use the additional “d” subscript to indicate the distance function.

$$\begin{cases} P_c(\tau) = \alpha_{c3}\tau^3 + \alpha_{c2}\tau^2 + \alpha_{c1}\tau + \alpha_{c0} \\ \alpha_{c0} = f(t_n) \\ \alpha_{c1} = \dot{f}(t_n)\Delta t \\ \alpha_{c2} = -3f(t_n) - 2\dot{f}(t_n)\Delta t + 3f(t_{n+1}) - \dot{f}(t_{n+1})\Delta t \\ \alpha_{c3} = 2f(t_n) + \dot{f}(t_n)\Delta t - 2f(t_{n+1}) + \dot{f}(t_{n+1})\Delta t \\ \Delta t = t_{n+1} - t_n \end{cases} \quad (4.55)$$

with τ uniformly spanning the interval ($0.0 < \tau < 1.0$).

If the derivative of the function is used, the cubic splining still applies, simply use the first and second derivatives instead of the function and first derivative. Remember that τ varies from 0.0 to 1.0. Now, extract the real, distinct root(s), τ_{droot} , of $P_{cd}(\tau)$ on the interval 0.0 to 1.0. This is necessary to distinguish the results from the ellipsoidal function we’ll examine shortly. If

$$\left. \frac{dP_{cd}(\tau)}{d\tau} \right|_{\tau=\tau_{droot}} > 0 \quad (4.56)$$

a local minimum range exists. The associated time of the close approach is

$$t_{CA}(\tau_{droot}) = t_n + \tau_{droot}\Delta t \quad (4.57)$$

where t_n is an endpoint of the time interval containing the minima, and $\Delta t = t_{n+1} - t_n$.

Once we have a candidate close approach, we must identify an error ellipsoid to model a “safety zone” around the primary object. We need this zone because propagation and differential correction techniques introduce errors into the orbit solution. Although we get a number indicating the object’s mean location at a point in time (propagation of the state), an uncertainty volume is actually defined by the propagation of the covariance matrix. The propagation of the covariance depicts the object’s probability of occupying a certain location within the overall region defined by the propagation of the state. The size of the covariance matrix is dependent on the original formation in the differential correction process, and the techniques used to move it through time. The combination of these features specifies a zone where the two objects are in close proximity. The length of time of this zone depends on the application. Typically, the propagation error is greatest in the along-track direction because time errors greatly displace the object along the orbital path. This tends to create an elongated zone (football shaped) about the primary object’s orbit. Three axes characterize the zone. We often set these values

arbitrarily to very large values because many systems don't have covariance information. For precise work, you should try to get accurate covariance estimates and use them when characterizing the error region. If you've done your differential-correction processing correctly, your error region should be fairly small. You want to create this region so the major axis is collinear with the primary object's velocity vector, \mathbf{v}_p . This is the NTW coordinate system.

It's important to obtain a function that determines the entry and exit conditions for the secondary object from the ellipsoidal region about the primary object. We're only interested in a two-dimensional result, so if we project \mathbf{r}_d along the primary axis (T) using the Pythagorean Theorem, we can use the general form of an ellipse (with a and b as the major and minor axes) to achieve the desired result.

$$\frac{T^2}{a^2} + \frac{N^2}{b^2} = 1, \quad \begin{cases} T = \frac{\mathbf{r}_d \cdot \mathbf{v}_p}{|\mathbf{v}_p|} \\ N^2 = \mathbf{r}_d \cdot \mathbf{r}_d - T^2 \end{cases} \quad (4.58)$$

Substituting T and N into the ellipse equation gives us the ellipsoidal function

$$f_e(\tau) = \frac{(\mathbf{r}_d \cdot \mathbf{v}_p)^2}{(\mathbf{v}_p \cdot \mathbf{v}_p)^2} + \frac{(\mathbf{r}_d \cdot \mathbf{r}_d) - \frac{(\mathbf{r}_d \cdot \mathbf{v}_p)^2}{(\mathbf{v}_p \cdot \mathbf{v}_p)}}{b^2} - 1 \quad (4.59)$$

where three conditions are possible. If $f_e(\tau) > 0$, the secondary is out of range. $f_e(\tau) = 0$ occurs as the secondary satellite enters or exits; the satellite is at the surface of the ellipsoidal region. $f_e(\tau) < 0$ means the secondary is within the specified range of the primary. We need the primary object's velocity only to orient the NTW coordinate system. It's not part of the overall logic. For a sphere, the minor and major axes are equal, so we don't need the projections, and the ellipsoidal function is

$$f_e(\tau) = \frac{\mathbf{r}_d \cdot \mathbf{r}_d}{a^2} - 1 \quad (4.60)$$

Let's examine the entry and exit points more closely. To determine the times of entry and exit, you must compute the coefficients α_{ce} from Eq. (4.61) for the cubic polynomial equation $P_{ce}(\tau)$. To do so, you need two endpoints p_1 , p_4 and two intermediate points, p_2 , p_3 for the interval.

$$\begin{cases} \alpha_{c0} = p_1 \\ \alpha_{c1} = \frac{(\tau_2^3 - \tau_2^2)(p_2 - p_1) + (\tau_1^2 - \tau_1^3)(p_3 - p_1) + (\tau_1^3 \tau_2^2 - \tau_1^2 \tau_2^3)(p_4 - p_1)}{DET} \\ \alpha_{c2} = \frac{(\tau_2 - \tau_2^3)(p_2 - p_1) + (\tau_1^3 - \tau_1)(p_3 - p_1) + (\tau_1 \tau_2^3 - \tau_1^3 \tau_2)(p_4 - p_1)}{DET} \\ \alpha_{c3} = \frac{(\tau_2^2 - \tau_2)(p_2 - p_1) + (\tau_1 - \tau_1^2)(p_3 - p_1) + (\tau_1^2 \tau_2 - \tau_1 \tau_2^2)(p_4 - p_1)}{DET} \\ DET = \tau_1^3 \tau_2^2 + \tau_1^2 \tau_2 + \tau_1 \tau_2^3 - \tau_1^3 \tau_2 - \tau_1^2 \tau_2^3 - \tau_1 \tau_2^2 \end{cases} \quad (4.61)$$

The endpoints are simply

$$\begin{cases} p_1 = f_e(0), & t = t_n \\ p_4 = f_e(1), & t = t_{n+1} \end{cases} \quad (4.62)$$

But you'll need to reconstruct the intermediate points from the distance functions of Eq. (4.52), assuming you can't just get the function values by invoking the propagator. This reconstruction involves determining the polynomial coefficients P_{ce} as before. Also, you must reconstruct the primary velocity and acceleration vectors at times t_n and t_{n+1} to produce the cubic coefficients α_{ce} . The intermediate points become

$$\begin{cases} p_2 = f_e(\tau_1), & \tau_1 = 1/3 \\ p_3 = f_e(\tau_2), & \tau_2 = 2/3 \end{cases} \quad (4.63)$$

Note that τ_1 and τ_2 are evenly spaced on the interval. If a minimum distance occurs in the interval, the values change to distribute the other root by halving the largest remaining interval.

By including the minimum distance, you'll force the $P_{ce}(\tau)$ curve to pass through that point; this ensures an entry or exit won't be missed due to ill-conditioned endpoints, as long as the quintic functions accurately represent relative distance components. Extract the real root(s), τ_{eroot} of $P_{ce}(\tau)$ on the interval 0.0 to 1.0.

Use a refining process to find the entry or exit time if a valid root exists. This step improves the accuracy of the $P_{ce}(\tau)$ curve in the region of τ_{eroot} by reconstructing the intermediate point p_2 at the root. If you find a second root, recompute p_3 ; otherwise, it remains associated with the minimum distance. Then, recompute new α_{ce} coefficients for the $P_{ce}(\tau)$ curve. Also recompute the real root(s), τ_{eroot} , of $P_{ce}(\tau)$ on the interval from 0.0 to 1.0, so the time of entry or exit becomes $\Delta t = t_{n+1} - t_n$.

$$t_{\text{EE}}(\tau_{\text{eroot}}) = t_n + \tau_{\text{eroot}}\Delta t, \quad \Delta t = t_{n+1} - t_n \quad (4.64)$$

An exit exists if

$$\left. \frac{dP_{ce}(\tau)}{d\tau} \right|_{\tau=\tau_{\text{eroot}}} > 0 \quad (4.65)$$

Otherwise, it's an entry.

The verification of A-N algorithm by the authors of this book finds it is significant to introduce the acceleration information into the algorithm. It is hard to get correct result if just take position and velocity information into the two-point,

three-order (cubic) polynomial interpolation. The two-point, five-order (quintic) polynomial interpolation using the position, velocity and acceleration information are required. This conclusion has been validated by the U.S. and Russian satellite's collision event. In addition, the relative distance function must be defined to be the square of the distance instead of the distance itself. Otherwise, the computation may be unstable due to the small denominators. If the objects' acceleration vectors aren't available, the vector can be approximated using the J_2 perturbative equation or two-body equation. The accuracy is adequate.

References

1. Zheng QY, Wu LD (2004) A computation method to warn the collision event between space probe and debris. *Acta Astronomica Sinica* 45(4):422–427
2. Liu J, Wang RL, Zhang HB (2004) Space debris collision prediction research. *Chin J Space Sci* 24(6):462–469
3. Hoots FR, Crawford LL, Roehrich RL (1984) An analytic method to determine future close approaches between satellites. *Celest Mech* 33(2):143–158
4. Qiu HX, Zhu ZM, Wu LD (2005) A study of collision analysis method for spacecraft. *J Astronaut* 26(3):257–261
5. Dybczynski PA, Jopek TJ, Serafin RA (1986) On the minimum distance between two Keplerian orbits with a common focus. *Celest Mech* 38:345–356
6. Kholshchevnikov KV, Vassiliev NN (1999) On the distance function between two Keplerian elliptic orbits. *Celest Mech Dyn Astron* 75:75–83
7. Gronchi GF (2005) An algebraic method to compute the critical points of the distance function between two Keplerian orbits. *Celest Mech Dyn Astron* 93:295–329
8. Gronchi GF, Tommei G (2007) On the uncertainty of the minimal distance between two confocal Keplerian orbits. *Discrete Continuous Dyn Syst Ser B* 7(4):755–778
9. Murison MA, Munteanu A (2006) On the distance function between two confocal Keplerian orbits. In: AIAA/ AAS astrodynamics specialist conference and exhibit, Keystone, Colorado
10. Armellini R, Di Lizia P, Berz M et al (2010) Computing the critical points of the distance function between two Keplerian orbits via rigorous global optimization. *Celest Mech Dyn Astron* 107:377–395
11. Hoots FR (1997) The future of artificial satellite theories, hybrid ephemeris compression model. *Celest Mech Dyn Astron* 66:51–60
12. Liou JC (2005) Accidental collisions of cataloged satellites identified. *Orbital Debris Q News* 9(2):1
13. Alfano S, Negron D (1993) Determining satellite close approach. *J Astronaut Sci* 41(2):217–225
14. Alfano S (1994) Determining satellite close approach Part II. *J Astronaut Sci* 42(2):143–152
15. Li J, Xiao YL (2007) An improved efficient algorithm for close approach determination of space objects. *Acta Aeronautica et Astronautica Sinica* 28(S):s42–s48
16. Rodriguez JRA, Martinez - Fadrique FM, Klinkrad H (2002) Collision risk assessment with a smart sieve method. In: Proceedings of joint ESA/NASA space-flight conference, Noordwijk, Netherland

17. Klinkrad H (2006) Space debris-models and risk analysis. Springer-Praxis, New York
18. Wang H, Huang H (2008) Methods to determine close approaches among orbiting objects. *J Astronaut* 29(6):1747–1751
19. Faulds AL, Spencer DB (2003) Satellite close-approach filtering using genetic algorithms. *Spacecraft Rockets* 40(2):248–252

Chapter 5

Calculation of Collision Probability

The conjunction assessment is a process as follow. Collision risk parameters (such as closest approach distance, collision probability, maximum collision probability) are obtained from collision risk assessment by using the predicted orbital states and covariance informations. The risk parameters are judged according to certain criteria whether they are in the danger-region.

The collision probability (P_c) between two objects in the conjunction is one of the most important criteria for conjunction risk assessment. The calculation of P_c , using information of position and velocity vectors and associated error covariance at the TCA, has been well studied. However, the existing methods did not reveal the direct and explicit connection between the P_c and the collision parameters. It is difficult to find the relationship between P_c and its influencing factors such as conjunction geometry and position error covariance without the connection. The relationship is important to decision maker for P_c sensitivity analysis. The sensitivity of P_c to covariance is significant for maximum P_c estimation, false alarm and miss alarm analysis, and probability dilution analysis.

This chapter derives the explicit expressions of collision probability (P_c) in terms of the components of relative position or the conjunction geometries in the cases of circular orbit and general orbit based on the analysis of the closest approach. The error of the explicit expression and the bounds of eccentricity to which the explicit expression is applicable are also determined.

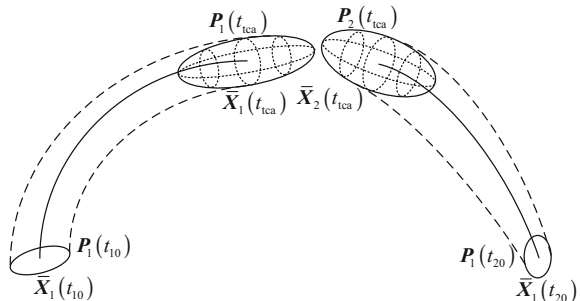
5.1 Calculation Method of Collision Probability and Simplification

5.1.1 Overview of Calculation Method of Collision Probability

Conjunction assessment and collision avoidance are progressively important with the continuous increase of on-orbit space objects. The closest approach distance and the collision probability (P_c) are two kinds of important criteria for conjunction risk assessment. The calculation of collision probability (P_c) is one of the fundamental techniques for the conjunction assessment and collision avoidance of space objects. The general description of the collision probability problem is follow. The estimated state (position and velocity) vector $\bar{\mathbf{X}}_1(t_{10})$ initial error covariance matrix $\mathbf{P}_1(t_{10})$ of the primary object are known at its epoch t_{10} ; the estimated state (position and velocity) vector $\bar{\mathbf{X}}_2(t_{20})$ initial error covariance matrix $\mathbf{P}_2(t_{20})$ of the secondary object are known at its epoch t_{20} . The state vector $\mathbf{X} = (\mathbf{r}, \mathbf{v})$ and the covariance matrix \mathbf{P} may be described in the same coordinate system (the ECI coordinate system for instance), or described in their own satellite-based coordinate system. The orbit (or state vector) and error covariance of each object are computed forward using orbit and covariance propagation models (analytical model or high-precision model). By an appropriate close approach analysis method, the time of closest approach (t_{tca}) between two objects and the corresponding state vectors and covariances $\bar{\mathbf{X}}_1(t_{tca})$, $\mathbf{P}_1(t_{tca})$, $\bar{\mathbf{X}}_2(t_{tca})$, $\mathbf{P}_2(t_{tca})$ associated with two objects can be obtained. The safety-radii of two objects are R_1 and R_1 . The P_c could be calculated using information of position and velocity vectors and associated error covariance ($\bar{\mathbf{X}}_1(t_{tca})$, $\mathbf{P}_1(t_{tca})$, $\bar{\mathbf{X}}_2(t_{tca})$, $\mathbf{P}_2(t_{tca})$) at the TCA. This process is illustrated in Fig. 5.1.

Because the perturbations such as the non-sphericity of the Earth and the atmospheric drag will cause substantial deviations from the Keplerian orbits, the propagation model of orbits and covariance must be accurate enough, otherwise the collision probability will be off.

Fig. 5.1 Description of the problem of P_c calculation



Currently, the main criteria is P_c between two objects in the conjunction, which is calculated using information of position and velocity vectors and associated error covariance at the TCA. The conjunction can be divided into the linear relative motion (short-term conjunction) case and the nonlinear relative motion (long-term conjunction) case in terms of the different conjunction geometries.

1. Linear relative motion (short-term conjunction)

The relative velocity at the point of closest approach is deemed sufficiently large to ensure a brief encounter time and static covariance for most conjunction events, where the angle between orbital planes is sufficiently large [1]. In these short-term conjunctions, the calculation of P_c is based on the assumption that the relative motion is linear and there is no velocity uncertainty, so that the position error ellipsoid during the encounter is constant and equal to the value at the estimated conjunction. The position uncertainties can be represented by three-dimensional Gaussian distributions. The collision probability is defined as the probability that the miss distance between two objects is less than the sum of their safety-radii.

Each object's positional uncertainties are combined and their radii summed at the point of closest approach. The resultant is projected into a plane perpendicular to the relative velocity, where the P_c is calculated as an integral of a two-dimensional (2D) Gaussian probability density function (PDF) over the circular collision cross-sectional area. There is no closed-form solution to this 2D integral problem. It is significant to find a computing method to the 2D integral with high precision and low computing burden.

Foster and Estes [2], Chan [1, 3–5], Alfano [6–8], Patera [9–11] have discussed the methods for computing P_c for short-term conjunction.

The method developed by Foster [2] evaluates the P_c numerically by examining the plane perpendicular to relative motion, dividing the combined object's circular cross-section in concentric circles and radial straight lines using polar coordinates in the conjunction plane [2]. Patera [9–11] of the Aerospace Corporation developed a mathematically equivalent model as a one-dimensional line integral. The probability density is symmetrized enabling the two-dimensional integral to be reduced to a one-dimensional path integral. Its evaluation is performed numerically by taking short line segments around a closed contour [9–12]. Alfano [6–8] of CSSI developed a series expression as a combination of error (erf) functions and exponential terms based on a one-dimensional PDF. It is numerically evaluated using well-known software already developed for error functions [6–8, 13]. Chan [3–5] of the Aerospace Corporation developed an analytical series expression based on representative, present-day collision probability parameters. He transforms the two-dimensional Gaussian PDF to a one-dimensional Rician PDF and uses the concept of equivalent areas. This model involves the evaluation of an analytical expression containing two exponential terms. This expression has the added benefit of being easily differentiated for other types of probability analysis such as maximum P_c analysis and sensitivity analysis [3–5, 14]. The algorithms of Foster, Patera and Alfano are essentially numerical, while Chan's method is analytical.

Alfano [15] compared the numerical results obtained by these algorithms with validation results of the 2D probability integral over a very wide range of collision parameters. The comparison revealed that Chan's method is the fastest but is also the most restrictive due to relative object size limitations. Patera's method produces good results, especially with object-oriented formulation. Alfano's method determines the number of integrations steps on a case-by-case basis. Foster's method is the slowest, but can be sped up by increasing the step size for many cases without adversely affecting accuracy.

The Recommended Standard for Conjunction Data Messages (CDMs) has been prepared by the Consultative Committee for Space Data Systems (CCSDS). The CDM described in this Recommended Standard is the baseline concept for conjunction information interchange applications between interested parties. This Recommended Standard establishes a common framework and provides a common basis for the format of conjunction information exchange between originators of conjunction assessment data and satellite owner/operators. The available options of method used to calculate the collision probability are FOSTER-1992, CHAN-1997, PATERA-2001, and, ALFANO-2005 [16].

In China, Wang Hua [17–22] of the National University of Defense Technology, Cheng Tao [23, 24], Feng Hao [25], Yang Xu [26] in the Liu Jing team of the National Space Science Center of the Chinese Academy of Sciences, Zhang Ming-xuan [27] of the Harbin Institute of Technology, Wu Bo [28] of the Information Engineering University have studied the calculation of collision probability.

2. Nonlinear relative motion (long-term conjunction)

The calculation of collision probability can be transformed from 3D integral to 2D integral in the conjunction plane under the assumption of linear relative motion. The conjunction is long-term if the relative velocity is small in the conjunction, both the magnitude and direction of the relative velocity will change during the conjunction. The relative becomes nonlinear. For example, satellites in the formation flight or space rendezvous, or the geosynchronous satellites have low relative velocity (meters per second or less) so that the time they spend in the encounter region is appreciable. The linear relative motion region might be smaller than the encounter region for very low relative velocity. Even if one propagates to the point of closest approach and applies current methods, an incorrect collision probability will be obtained.

The calculation of collision probability is more complex in the nonlinear relative motion than in the linear one. First, differ from the one in linear relative motion, the collision probability is changeable during the nonlinear encounter. Second, the problem can not be transformed to a 2D integration, but requires a numerical integration of the probability density through the volume swept out by the combined hard body of the space vehicles. This integration is complicated by the changing orientation of the hard body and combined position-error ellipsoid throughout the encounter. In addition, integration limits on this volume integral complicate the problem.

Chan investigated when it is justified to assume rectilinear relative motion so as to replace a three-dimensional integral by a two-dimensional integral for computing the collision probability when two spacecraft are in close proximity [1]. Chan also concerned with the modeling of long-term spacecraft encounters for the purpose of computing collision probability [29]. Patera presented a method for calculating the collision probability between two space vehicles when the relative motion is nonlinear, which is developed using contour integration methodology. The method involves transforming the problem to a scaled frame in which the error covariance matrix is symmetric in three dimensions [30, 31]. Slater examined the evolution of the collision probability for satellites in formation flight and determined an effective maneuver algorithm that can minimize the probability of collision while reducing the energy expenditure in the maneuver [32]. Alfano pointed out that the dimension associated with relative velocity must be reintroduced for nonlinear motion. This can be simply done by breaking the collision tube into sufficiently small cylinders such that the sectional motion is nearly linear, computing the linear probability associated with each section, and then summing. An alternate method is that creates a voxel grid in Mahalanobis space, computes the probability of each affected voxel as the combined object passes through the space, and sums [33]. McKinley developed an algorithm that reduces the complex nonlinear relative motion into small linear segments that are easy to evaluate [34]. Xu Xiao-li of the Purple Mountain Observatory of Chinese Academy of Sciences studied the method used to calculate collision probability for nonlinear relative motion based on McKinley's method [35]. Zhang Ge of Dalian University of Technology studied the prediction of collision probability and collision avoidance for satellites in formation flight [36]. Luo Ya-zhong and Liang Li-bo of National University of Defense Technology investigated two quantitative performance indexes of space rendezvous trajectory safety in consideration of navigation errors and control errors [37–41].

3. Explicit expression of collision probability

These methods previous introduced did not reveal the direct and explicit connection between the P_c and the collision parameters. It is difficult to find the relationship between P_c and its influencing factors such as conjunction geometry and position error covariance without the connection. The relationship is important to decision maker for P_c sensitivity analysis. Given that the parameters needed for computing the P_c are uncertain, one should be cautious in making a decision based upon P_c computed using nominal values. It would be prudent to determine the sensitivity of the computed probability to expected variations in the parameters describing the computation [42]. The sensitivity analysis of P_c is to investigate the effects of influencing factors (conjunction geometry, error covariance, object's size) on P_c calculations. The sensitivity of P_c to conjunction geometry is significant for optimal strategy making of avoidance maneuver and analyzing the effect of bias of orbital prediction on P_c calculation. The P_c calculation is most sensitive to the quality of the covariance. The improper covariance will trigger false alarm or miss

alarm. The sensitivity of P_c to covariance is significant for maximum P_c estimation, false alarm and miss alarm analysis, and probability dilution analysis.

The closest approach distance (and its components or relative positions in object-centered coordinates system) and the P_c are two kinds of important criteria for conjunction risk assessment, in accordance with the Box method and the P_c method. The P_c is a comprehensive risk assessment criterion with consideration of conjunction geometries, position error covariance, and sizes of space objects. Meanwhile, the miss distance, relative positions, relative velocity, and the encounter angle are also useful information for assessing collision risk. The P_c calculation is a point estimate of the risk based on the state and covariance information that were utilized. As the criteria of conjunction risk, the P_c is more sophisticated than relative distance, but is susceptible to uncertainty of covariance. In some cases where the error covariance are not available, the distance-based Box method is more reliable. Since the relative distance thresholds are useful in conjunction events' screening process, it is significant to understand the relationship between probability threshold and distance threshold.

In practical conjunction assessment, these two criteria are often provided together. For example, [43] pointed out that United States Joint Space Operations Center (JSpOC)'s Conjunction Summary Message (CSM) contains the time of closest approach (TCA), miss distance, relative speed, closest approach relative position (radial, intrack, cross-track), relative velocity (radial, in-track, crosstrack) et al. The Consultative Committee for Space Data Systems (CCSDS)'s Recommended Standard for Conjunction Data Message (CDM) includes the TCA, miss distance, relative speed, radial, transverse, and normal components of relative position/velocity in the primary RSW frame, the collision probability, the computing method of collision probability.

Alfriend et al. presented an approach for determining an upper limit of P_c and a parametric analysis to determine the sensitivity of P_c to the encounter geometry and covariance [44]. In order to gain insight into the primary factors that affect P_c a simplified example was developed. The error covariance was assumed to be aligned with the radial, in-track and out-of-plane directions. It was also assumed that the objects are in near circular orbit so that their velocities are equal. The miss vector at conjunction can be resolved into two components: radial separation distance and horizontal plane separation distance. The horizontal plane separation distance can be obtained using the distance of objects from the intersection of the two orbital planes and the angle between two orbital planes. Based on this analysis, they presented a collision probability integral expression in terms of radial distance, horizontal distance, angle between orbital planes, and scaling factor of error covariance. An approximate value of P_c can be calculated by assuming the probability density is constant over the collision circle.

Chan [3] constructed two planes perpendicular to the vector defined by the cross product of velocities of primary and secondary objects: one containing the primary velocity vector and the other containing the secondary velocity vector [3, 14]. Then he projected secondary velocity onto the primary velocity plane. By analyzing the minimum separation at conjunction, the minimum separation between two objects

can be resolved into the distance (height) between two projection planes and the minimum separation distance in the projection plane. Based on the “general simplified model” that the covariance ellipsoids are aligned with the velocity coordinate axes for both the objects, the angle between the vector of minimum separation and the major axis of the covariance ellipse in the conjunction plane was given.

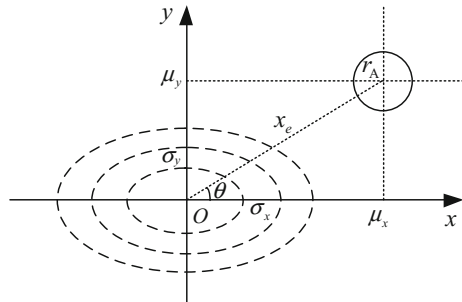
Both Alfriend et al. [44] and Chan [3] have resolved the miss distance at conjunction into two components. The Alfriend’s radial separation distance and horizontal plane separation distance are essentially the Chan’s height between two projection planes and the minimum separation distance in the projection plane in the special case of near circular orbits. The computation of miss distance’s components were not given by Alfriend et al. [44], while Chan [3] has discussed the details for computing the components from initial condition. The processes and results are somewhat complicated that will restrict its actual applications. Moreover, Alfriend et al.’s [44] expression of collision probability in terms of radial and horizontal distance was still in the form of one-dimensional integral with error functions (erf). Chan [3] has not considered this characteristic in his succeeding collision probability analysis. The relationship between the components of miss distance and the collision probability was not revealed explicitly in their works [3].

5.1.2 Simplification of Calculation Method of Collision Probability

Many previous studies on collision probability have revealed that the calculation of P_c in short-term conjunctions can eventually be transformed to an integral of two-dimensional anisotropic probability density function (PDF) over a circular collision cross-section area on the conjunction plane. On the conjunction plane, the circular collision cross-section area is centered at (μ_x, μ_y) , whose radius is r_A . The combined standard deviations are σ_x and σ_y . As illustrated in Fig. 5.2, the probability integral parameters on the conjunction plane are μ_x , μ_y , σ_x , σ_y , and r_A .

The P_c equals to the integration of PDF over the collision cross-section area, as seen in Eq. (5.1).

Fig. 5.2 Parameters of P_c density function and radius of collision cross-section area on the conjunction plane



$$P_c = \iint_{(x-\mu_x)^2 + (y-\mu_y)^2 \leq r_A^2} \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] dx dy \quad (5.1)$$

Chan [3, 5] presented an improved analytical expression for computing the P_c . On the conjunction plane, a scale transformation converts the two-dimensional Gaussian PDF to an isotropic Gaussian PDF and the circular cross-section to an elliptical one. By approximating this ellipse with an equivalent circle, the two-dimensional Gaussian distribution is replaced by a one-dimensional Rician distribution, the integral of which is known in the form of a convergent infinite series. His study shows that it is permissible to truncate after the first or second term and still obtain results to within 0.4% error for a wide range of collision parameters considered.

The authors of this book did analogous work independently in Refs. [45] and [46]. Differ from Chan who gave a sum of infinite terms to describe the infinite series, we presented the first term and the recursive expression of the infinite series, which are advantageous for programming of the method. The results will be given here.

Let the dimensionless parameters v and u as

$$\begin{cases} v = \frac{1}{2}\left(\frac{\mu_x^2}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2}\right) \\ u = \frac{r_A^2}{2\sigma_x\sigma_y} \end{cases} \quad (5.2)$$

The first term of the infinite series is

$$P_0 = e^{-v}(1 - e^{-u}) \quad (5.3)$$

The recursive relationship of the k th term P_k and the $(k-1)$ th term P_{k-1} is

$$P_k = \frac{v}{k}P_{k-1} - \frac{u^k v^k}{k!k!} e^{-(v+u)}, \quad k \geq 1 \quad (5.4)$$

Let

$$a_k = \frac{v}{k}, \quad b_k = \frac{u^k v^k}{k!k!} e^{-(v+u)}, \quad k \geq 1 \quad (5.5)$$

Then the recursive expression will be

$$P_k = a_k P_{k-1} - b_k, \quad k \geq 1 \quad (5.6)$$

Any term $P_k (k \geq 1)$ could be calculated based on the first term and the recursive expression. If only the first $n+1$ terms in the series are retained as the approximation of the collision probability in the conjunction assessment

$$P_c \approx P_0 + P_1 + \cdots + P_n = \sum_{k=0}^n P_k \quad (5.7)$$

Then the truncation error S_n satisfies the inequality

$$S_n < \frac{1}{n!(n+1)!} u^{n+1} v^n e^{-v} e^{uv} \quad (5.8)$$

Reference [46] indicates that if the first term P_0 is taken as the approximation of the probability integral, the relative truncation error S_1/P is of the order of magnitude of 10^{-5} or smaller; if the first two terms $P = P_0 + P_1$ are taken, the relative truncation error S_2/P is as small as 10^{-9} . Therefore, if only the first term is retained the relative truncation error is neglectable. It turns out that for most space object's encounters this truncation after the first term yields an error that appears in the third significant digit. Thus, it suffices to retain only the first term Eq. (5.3). Substituting v and u in Eq. (5.2) into Eq. (5.3), we can obtain the approximate analytical expression of P_c

$$P_c = \exp \left[-\frac{1}{2} \left(\frac{\mu_x^2}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2} \right) \right] \left[1 - \exp \left(-\frac{r_A^2}{2\sigma_x\sigma_y} \right) \right] \quad (5.9)$$

One may need to retain more terms if the cross-sectional radius and the miss distance are approximately greater than the standard deviations. In addition, Chan made an approximation in representing the elliptical cross section by a circular cross section in the deduction. This second approximation ultimately determines the accuracy of using the Rician PDF in this modeling.

The approximate analytical expression Eq. (5.9) will service as the basic formula to analyze the collision probability.

5.2 Explicit Expression of Collision Probability in Circular Orbit

Most of space objects are in the circular or near-circular orbits, such as Low Earth Orbit (LEO), Middle Earth Orbit (MEO), and Geostationary Earth Orbit (GEO). As a consequence of the characteristic of orbit motion, the close approach or collision between two space objects always occur in the vicinity of points of closest approach between their orbits, as seen in Fig. 5.3.

In the case of circular orbits, the points of closest approach between orbits are along the line of intersection of orbital planes. For the objects in the near-circular orbit as illustrated in Fig. 5.4, the following assumptions are made:

Fig. 5.3 Conjunction always occurs in the vicinity of points of closest approach between orbits

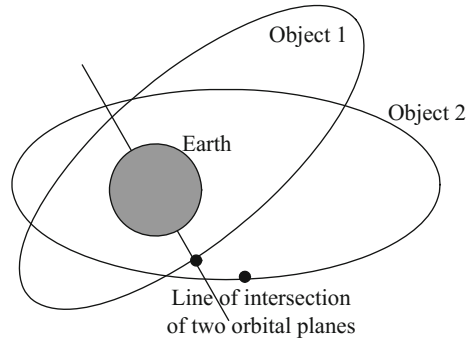
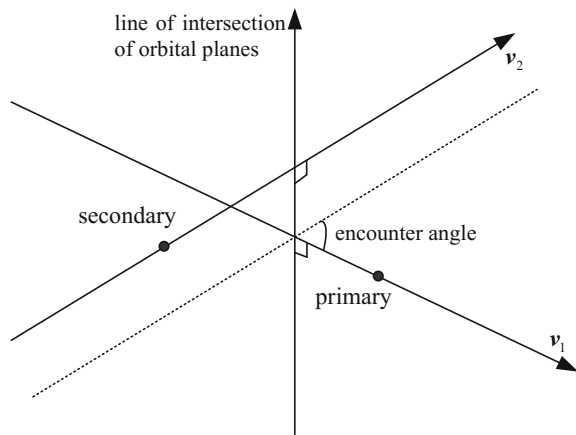


Fig. 5.4 Circular orbit assumption and the conjunction configuration



- (1) Two space objects involving conjunction are both in circular orbits; their velocity vectors are perpendicular to their position vectors.
- (2) The altitudes of two near-circular orbits are essentially same, so two objects' velocities have essentially the equal magnitude.
- (3) The angle between two velocity vectors (encounter angle) is equal to the angle between two orbital planes.
- (4) The error covariance is aligned with the radial, in-track and cross-track directions, there are no off-diagonal elements for covariance in the RSW frame.

This section will derive the explicit expressions of collision probability (P_c) in terms of the RSW components of relative position or the conjunction geometries (crossing altitude difference and time difference of the line of intersection of two orbital planes, the angle between orbital planes, etc.) in the case of circular orbit.

5.2.1 Conjunction Geometry and Distance Analysis

The relative position and velocity relationship between two space objects at TCA could be represented by either the components of relative position and velocity or the conjunction geometries such as the crossing altitude difference and time difference of the line of intersection of the two orbital planes. Accordingly, the collision probability can also be represented by either the closest approach distance and its components or the conjunction geometries. The two methods will be discussed below.

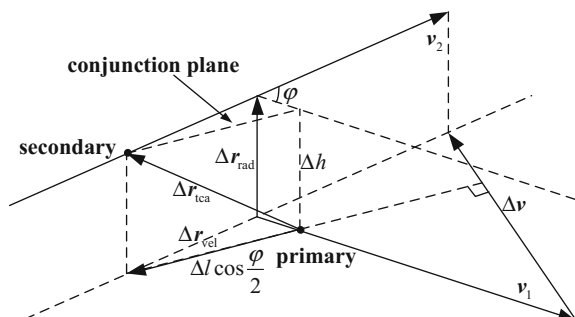
1. Analysis of closest approach distance

The closest approach distance between two orbits is in nature the closest approach distance between two objects without considering the synchronization. It is obvious from differentiable curves theory that the necessary condition of closest approach between two orbits is that the relative position vector is perpendicular to both objects' velocity vectors, i.e., the relative position vector is parallel to the common perpendicular line to two velocities. The common perpendicular line is identical to the line of intersection of orbital planes under the circular orbit assumption. It is also obvious that the necessary condition of closest approach between two objects moving around their orbits is that the relative position vector Δr_{tca} between two objects is perpendicular to their relative velocity vector Δv and on the conjunction plane at the time of closest approach. The conjunction plane is perpendicular to the relative velocity vector Δv .

Figure 5.5 illustrates the closest approach between orbits and the closest approach between objects. The plane contains the dashed rectangle in Fig. 5.5 is the conjunction plane which is perpendicular to the relative velocity. Two objects are at the diagonal vertices of this rectangle at TCA.

Therefore, as shown in Fig. 5.5, the relative position vector at the TCA, Δr_{tca} , can be resolved into two components: (1) closest approach vector between orbits, Δr_{rad} , which is in the direction of common perpendicular line to two objects' velocity vectors, in the circular orbit case it is in the radial direction; and (2) closest approach vector in the plane perpendicular to the common perpendicular line to two velocities (referred to as the horizontal plane hereafter), Δr_{vel} .

Fig. 5.5 The closest approach vector between orbits Δr_{rad} and the closest approach vector at TCA Δr_{tca}



$$\Delta \mathbf{r}_{\text{tca}} = \Delta \mathbf{r}_{\text{rad}} + \Delta \mathbf{r}_{\text{vel}} \quad (5.10)$$

where $\Delta \mathbf{r}_{\text{rad}}$ and $\Delta \mathbf{r}_{\text{vel}}$ are perpendicular to each other. Let Δr_{tca} , Δr_{rad} , and Δr_{vel} are respectively the magnitudes of three vectors, the miss distance at the TCA can be given by

$$\Delta r_{\text{tca}}^2 = \Delta r_{\text{rad}}^2 + \Delta r_{\text{vel}}^2 \quad (5.11)$$

We can define the encounter coordinate system, $o-xyz$ based on the analysis of closest approach distance. The origin o is at the centroid of primary object, the z -axis points in the direction of relative velocity $\Delta \mathbf{v}$, the x -axis points in the direction of closest approach vector between orbits (i.e. the line of intersection of orbital plane, it is in the radial direction when two orbits are circular) $\Delta \mathbf{r}_{\text{rad}}$, the y -axis lies in the horizontal plane and perpendicular to the relative velocity vector. The encounter coordinate system has a right-handed sense. The x - y plane perpendicular to the relative velocity is the conjunction plane, the closest approach vector $\Delta \mathbf{r}_{\text{tca}}$ is in this plane, as shown in Fig. 5.6.

At the TCA, the coordinates of distributing center of the position error ellipsoid of secondary object in the encounter coordinate system are

$$\begin{cases} \mu_x = \Delta r_{\text{rad}} \\ \mu_y = \Delta r_{\text{vel}} \end{cases} \quad (5.12)$$

We can also define the RSW object-centered coordinate system of primary object. The origin is at the centroid of primary object. The R axis (radial) always points from the Earth's center along the radius vector toward the satellite as it moves through the orbit, the S axis (in-track) points in the direction of (but not necessarily parallel to) the velocity vector and is perpendicular to the radius vector, the W axis (cross-track) is normal to the orbital plane. The S axis is aligned with the velocity vector for circular orbit, as shown in Fig. 5.7.

Fig. 5.6 Definition of the encounter coordinate system

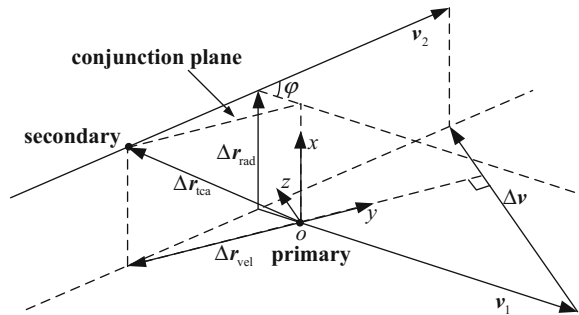


Fig. 5.7 Definition of the RSW coordinate system of primary object

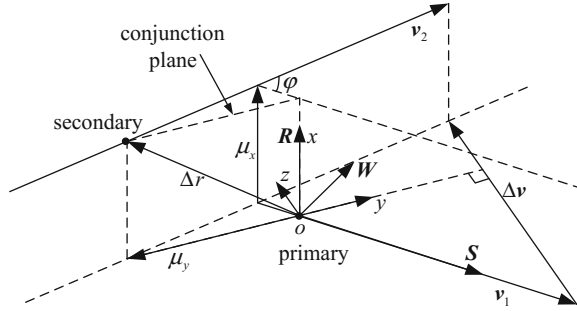
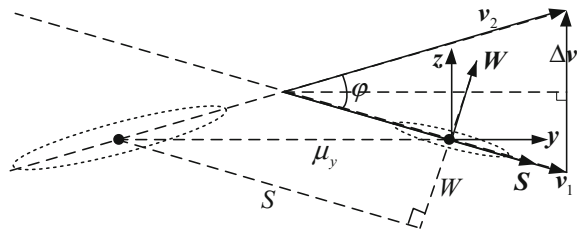


Fig. 5.8 Closest approach distance in the horizontal plane



At the TCA, the components of relative position in the RSW coordinate system are R , S , and W , the miss distance $\Delta r_{\text{tca}}^2 = R^2 + S^2 + W^2$. It is obvious from Fig. 5.7 that the R -axis of RSW coordinate system is identical to the x -axis of encounter coordinate system, so the x -coordinate of secondary object in the encounter coordinate system is R . The closest approach distance in the horizontal plane is determined by S and W coordinates, as shown in Fig. 5.8. So the coordinates of secondary object in the encounter coordinate system can be expressed using the components of relative position

$$\begin{cases} \mu_x = R \\ \mu_y = \sqrt{S^2 + W^2} \end{cases} \quad (5.13)$$

2. Conjunction geometries

The crossing altitude difference and time difference of the line of intersection of the two orbital planes are commonly used to pre-filter in the close approach analysis [47–49]. The conjunction geometries includes crossing altitude difference and time difference of the line of intersection of the two orbital planes, the angle between two orbital planes φ , and so on, as illustrated in Fig. 5.5.

The close approach between two space objects occurs in the vicinity of points of closest approach between their orbital planes. The altitude difference is Δh , and the distance between secondary object and the intersection line is Δl when the primary object arrives at the intersection line. This distance is due to the time difference Δt

of two objects crossing the intersection line, $\Delta l = v \cdot \Delta t$. The circular orbit assumption implies that the velocity vectors v_1 and v_2 are both perpendicular to the intersection line vector r , so is the relative velocity vector Δv , as shown in Fig. 5.8. The two sides of the dashed rectangle on the conjunction plane in Fig. 5.5 are Δh and $\Delta l \cos(\varphi/2)$, respectively. The closest approach distance equals to the length of the diagonal of the rectangle, i.e.

$$\Delta r_{\text{ca}} = \sqrt{\Delta h^2 + \Delta l^2 \cos^2(\varphi/2)} = \sqrt{\Delta h^2 + v^2 \cos^2(\varphi/2) \Delta t^2} \quad (5.14)$$

So the coordinates of secondary object in the encounter coordinate system can be expressed using the conjunction geometries

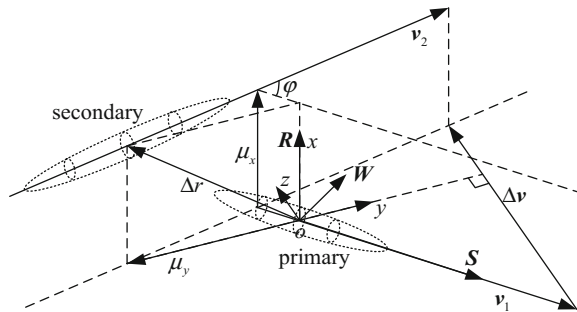
$$\begin{cases} \mu_x = \Delta h \\ \mu_y = \Delta l \cos \frac{\varphi}{2} = v \cos(\varphi/2) \Delta t \end{cases} \quad (5.15)$$

Equations (5.13) and (5.15) express the probability integral parameters (μ_x, μ_y) in terms of the RSW components of closest approach distance and the conjunction geometries, respectively.

5.2.2 Transformation and Projection of Error Covariance

The position error covariance matrices associated with the objects in near-circular orbit are generally defined in their object-centered RSW coordinate system, as illustrated in Fig. 5.9. Alfriend et al. [44] assumed that the error covariance is aligned with the radial, in-track and crosstrack directions, the radial and cross-track components are assumed equal. Chan [3] pointed out that in most cases for almost-circular orbits, the velocities of spacecraft or orbiting objects are essentially in the direction of the major principal axes. He assumed that the major axis of the covariance ellipsoid is aligned with the velocity vector. This is reasonable because much of the positional uncertainty is along this direction. He also noted that the intermediate axis is essentially (but not exactly) aligned in the cross-track direction

Fig. 5.9 Position error covariance ellipsoids of two objects which are aligned with the radial, in-track and cross-track directions



and the minor axis is essentially (but not exactly) aligned in the radial direction. These assumptions about the alignment of axes greatly simplify the analysis and still retain the essential features of the position error covariance.

Based on Alfriend and Chan's assumption, the covariance matrices of two objects in their RSW coordinate systems are given by

$$\begin{cases} \mathbf{P}_{1\text{RSW}} = \begin{bmatrix} \sigma_{1\text{R}}^2 & 0 & 0 \\ 0 & \sigma_{1\text{S}}^2 & 0 \\ 0 & 0 & \sigma_{1\text{W}}^2 \end{bmatrix} \\ \mathbf{P}_{2\text{RSW}} = \begin{bmatrix} \sigma_{2\text{R}}^2 & 0 & 0 \\ 0 & \sigma_{2\text{S}}^2 & 0 \\ 0 & 0 & \sigma_{2\text{W}}^2 \end{bmatrix} \end{cases} \quad (5.16)$$

where $\sigma_{1\text{R}}^2$, $\sigma_{1\text{S}}^2$, and $\sigma_{1\text{W}}^2$ are respectively the radial, in-track, and cross-track variances of primary object, $\sigma_{2\text{R}}^2$, $\sigma_{2\text{S}}^2$, and $\sigma_{2\text{W}}^2$ are of secondary object. In order to transform position error covariance to the encounter coordinate system, we must know the transformation matrices from the RSW system to the encounter system. The R -axis of RSW system is aligned with the x -axis of encounter system, hence we just need to rotate the RSW systems about their R -axis to get the encounter system. The required angles of rotation are $\varphi/2$ and $-\varphi/2$ respectively (see Fig. 5.8), the transformation matrices are given in Eq. (5.17).

$$\mathbf{M}_{1\text{RSW} \rightarrow \text{xyz}} = \mathbf{M}_1\left(\frac{\varphi}{2}\right), \quad \mathbf{M}_{2\text{RSW} \rightarrow \text{xyz}} = \mathbf{M}_1\left(-\frac{\varphi}{2}\right) \quad (5.17)$$

where $\mathbf{M}_1(\cdot)$ is the elementary rotation matrix describing rotation around the first axes.

Transforming $\mathbf{P}_{1\text{RSW}}$ and $\mathbf{P}_{2\text{RSW}}$ to the encounter system by using transformation matrices in Eq. (5.17), we can obtain

$$\begin{aligned} \mathbf{P}_1 &= \mathbf{M}_{1\text{RSW} \rightarrow \text{xyz}} \mathbf{P}_{1\text{RSW}} \mathbf{M}_{1\text{RSW} \rightarrow \text{xyz}}^T \\ &= \begin{bmatrix} \sigma_{1\text{R}}^2 & 0 & 0 \\ 0 & \sigma_{1\text{S}}^2 \cos^2 \frac{\varphi}{2} + \sigma_{1\text{W}}^2 \sin^2 \frac{\varphi}{2} & \frac{\sigma_{1\text{W}}^2 - \sigma_{1\text{S}}^2}{2} \sin \varphi \\ 0 & \frac{\sigma_{1\text{W}}^2 - \sigma_{1\text{S}}^2}{2} \sin \varphi & \sigma_{1\text{S}}^2 \sin^2 \frac{\varphi}{2} + \sigma_{1\text{W}}^2 \cos^2 \frac{\varphi}{2} \end{bmatrix} \end{aligned} \quad (5.18)$$

$$\begin{aligned} \mathbf{P}_2 &= \mathbf{M}_{2\text{RSW} \rightarrow \text{xyz}} \mathbf{P}_{2\text{RSW}} \mathbf{M}_{2\text{RSW} \rightarrow \text{xyz}}^T \\ &= \begin{bmatrix} \sigma_{2\text{R}}^2 & 0 & 0 \\ 0 & \sigma_{2\text{S}}^2 \cos^2 \frac{\varphi}{2} + \sigma_{2\text{W}}^2 \sin^2 \frac{\varphi}{2} & \frac{\sigma_{2\text{S}}^2 - \sigma_{2\text{W}}^2}{2} \sin \varphi \\ 0 & \frac{\sigma_{2\text{S}}^2 - \sigma_{2\text{W}}^2}{2} \sin \varphi & \sigma_{2\text{S}}^2 \sin^2 \frac{\varphi}{2} + \sigma_{2\text{W}}^2 \cos^2 \frac{\varphi}{2} \end{bmatrix} \end{aligned} \quad (5.19)$$

Under the assumption that the covariance is uncorrelated, the combined covariance is given by the sum of two individual covariances.

$$\mathbf{P}^* = \begin{bmatrix} \sigma_{1R}^2 + \sigma_{2R}^2 & 0 & 0 \\ 0 & (\sigma_{1S}^2 + \sigma_{2S}^2) \cos^2 \frac{\varphi}{2} + (\sigma_{1W}^2 + \sigma_{2W}^2) \sin^2 \frac{\varphi}{2} & \frac{1}{2} (\sigma_{1W}^2 - \sigma_{1S}^2 + \sigma_{2S}^2 - \sigma_{2W}^2) \sin \varphi \\ 0 & \frac{1}{2} (\sigma_{1W}^2 - \sigma_{1S}^2 + \sigma_{2S}^2 - \sigma_{2W}^2) \sin \varphi & (\sigma_{1S}^2 + \sigma_{2S}^2) \sin^2 \frac{\varphi}{2} + (\sigma_{1W}^2 + \sigma_{2W}^2) \cos^2 \frac{\varphi}{2} \end{bmatrix} \quad (5.20)$$

We just take two components x and y on the conjunction plane into account, and sum-up two covariance matrices to get two-dimensional combined error covariance matrix \mathbf{P} in conjunction plane associated with the relative position

$$\mathbf{P} = \begin{bmatrix} \sigma_{1R}^2 + \sigma_{2R}^2 & 0 \\ 0 & (\sigma_{1S}^2 + \sigma_{2S}^2) \cos^2 \frac{\varphi}{2} + (\sigma_{1W}^2 + \sigma_{2W}^2) \sin^2 \frac{\varphi}{2} \end{bmatrix} \quad (5.21)$$

Equation (5.21) indicates that the combined covariance matrix \mathbf{P} on the conjunction plane is precisely diagonal for all values of the encounter angle. Let the combined error variances of R, S, W directions of two objects be

$$\sigma_R^2 = \sigma_{1R}^2 + \sigma_{2R}^2, \quad \sigma_S^2 = \sigma_{1S}^2 + \sigma_{2S}^2, \quad \sigma_W^2 = \sigma_{1W}^2 + \sigma_{2W}^2 \quad (5.22)$$

The position error in R direction is not coupled with position error in S and W direction; one can directly sum up error variances of two objects in R direction to obtain the combined error variance σ_R^2 . Whereas position errors in S and W direction of two objects are coupled by the angle between their orbital plane. We define combined error variance in the horizontal plane as

$$\sigma_{SW}^2 = \sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2} \quad (5.23)$$

The combined error variances in the x and y directions on the conjunction plane are

$$\begin{cases} \sigma_x^2 = \sigma_R^2 = \sigma_{1R}^2 + \sigma_{2R}^2 \\ \sigma_y^2 = \sigma_{SW}^2 = \sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2} \end{cases} \quad (5.24)$$

We will discuss some of the properties associated with the combined error variance in the horizontal plane. Two objects are headed directly toward each other when $\varphi = \pi$, the combined error standard deviation is $\sigma_{SW} = \sigma_W$. Two objects are approaching each other broadside when $\varphi = \pi/2$, the combined error standard deviation $\sigma_{SW} = \sqrt{(\sigma_S^2 + \sigma_W^2)}/2$. Two objects are headed almost in the same direction when φ is small ($\varphi \rightarrow 0$), the combined error standard deviation $\sigma_{SW} = \sigma_S$.

Figure 5.10 illustrates the combined standard deviation σ_{SW} computed using Eq. (5.23) for various angles φ in two cases: (1) $\sigma_S = 100$, $\sigma_W = 20$ (line with marker “o”), and (2) $\sigma_S = 20$, $\sigma_W = 100$ (line with marker “Δ”). Figure 5.10 indicates that the combined standard deviation σ_{SW} will decrease with the increase

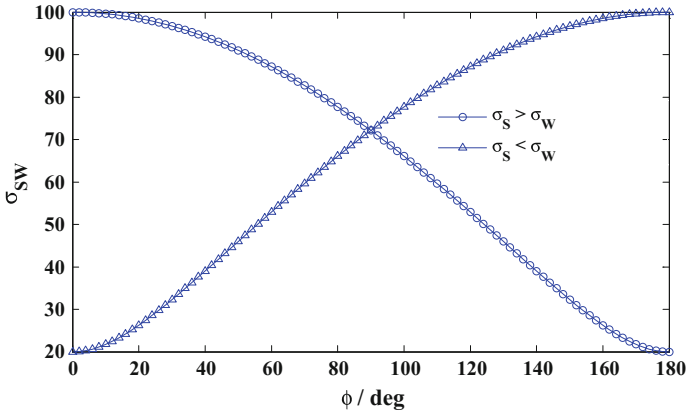


Fig. 5.10 Combined error variance in the horizontal plane σ_{SW} versus the angle between two velocity vectors when $\sigma_S > \sigma_W$ and $\sigma_S < \sigma_W$

of the angle φ when $\sigma_S > \sigma_W$, which is the ordinary situation. The combined standard deviation σ_{SW} will increase with the increase of the angle φ in some infrequent case when $\sigma_S < \sigma_W$.

So far, we have obtained explicit expression of probability integral parameters $\mu_x, \mu_y, \sigma_x,$ and σ_y from Eqs. (5.13), (5.15) and (5.24).

5.2.3 Explicit Expression of Collision Probability

The explicit expression of collision probability can be represented by either the components of the closest approach distance or the conjunction geometries.

Substituting Eqs. (5.13) and (5.24) to Eq. (5.2), we can express the dimensionless parameters v and u as

$$v = \frac{1}{2} \left(\frac{R^2}{\sigma_R^2} + \frac{S^2 + W^2}{\sigma_{SW}^2} \right), \quad u = \frac{r_A^2}{2\sigma_R\sigma_{SW}} \tag{5.25}$$

Substituting Eq. (5.25) into Eq. (5.9), we can obtain the explicitly expression of collision probability in terms of the components of the relative position at the TCA and combined error variances in their RSW coordinate systems.

$$P_c = \exp \left[-\frac{1}{2} \left(\frac{R^2}{\sigma_R^2} + \frac{S^2 + W^2}{\sigma_{SW}^2} \right) \right] \cdot \left[1 - \exp \left(-\frac{r_A^2}{2\sigma_R\sigma_{SW}} \right) \right] \tag{5.26}$$

where R, S, W are the radial, in-track, and cross-track components of relative position in primary object’s RSW coordinate system at the TCA; φ is the angle

between two velocity vectors; σ_R , σ_S , σ_W are combined position error standard deviations of two objects.

Substituting Eqs. (5.15) and (5.24) into Eq. (5.9), we can obtain the explicitly expression of collision probability in terms of the crossing altitude difference and time difference of the line of intersection of two orbital planes and combined error variances in their RSW coordinate systems.

$$P_c = \exp \left[-\frac{1}{2} \left(\frac{\Delta h^2}{\sigma_R^2} + \frac{v^2 \cos^2 \frac{\varphi}{2} \Delta t^2}{\sigma_{SW}^2} \right) \right] \left[1 - \exp \left(-\frac{r_A^2}{2\sigma_R \sigma_{SW}} \right) \right] \quad (5.27)$$

where Δh and Δt are the altitude difference and time difference; v is the magnitude of velocity of space objects; φ is the angle between two orbital planes; σ_R , σ_S , σ_W are combined position error standard deviations of two objects.

Equations (5.26) and (5.27) indicate that the explicit expression of collision probability includes two distance-related terms: the radial term R^2/σ_R^2 (or the altitude-related term $\Delta h^2/\sigma_R^2$) and the horizontal term (or the time-related term)

$$\frac{S^2 + W^2}{\sigma_{SW}^2} = \frac{S^2 + W^2}{\sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2}}, \quad \frac{v^2 \cos^2 \frac{\varphi}{2} \Delta t^2}{\sigma_{SW}^2} = \frac{v^2 \cos^2 \frac{\varphi}{2} \Delta t^2}{\sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2}} \quad (5.28)$$

From the collision probability's explicit expression Eqs. (5.26) and (5.27) we know that:

- (1) The error variances of each object in each direction are simply summed up to form the combined error variances, which simplifies the transformation and projection of the error and is convenient for the analysis of the problem.
- (2) The position errors in the horizontal plane (S and W) are coupled by the angle φ between their orbital planes. The horizontal combined variances are related to the horizontal components of relative position at conjunction (or the time difference Δt crossing the orbital intersection line). The position errors in the radial direction (R) are decoupled with the position error in the horizontal plane, we can directly sum up error variances of two objects in the R direction to obtain the combined error variance σ_R^2 . The radial combined variances are related to the radial component of the relative position at conjunction (or the altitude difference Δh crossing the orbital intersection line).
- (3) The P_c can be considered as a function of two mutually independent error variances, σ_R^2 and σ_{SW}^2 . The two error variances are related to the radial relative distance and the horizontal relative distance, respectively. For a non-coplanar conjunction the typical avoidance maneuver is a posigrade maneuver to increase in-track or radial separation at TCA. An operator could deduce by separating the contributions of R and SW directions that just radial error variance need to be considered when discuss the radial separation and its impact on collision risk. In similar fashion, the combined error in the horizontal plane should be considered when discuss in-track and cross-track separation.

The two forms of the explicit expression of P_c have their own advantages and applications. The distance components-represented explicit expression is useful in the analysis of the relationship between two risk assessment criteria (miss distance and collision probability). Since increasing the altitude difference and time difference crossing the intersection line of the orbital planes are principal strategy for collision avoidance maneuver of spacecrafts, the conjunction geometries-represented explicit expression is significant to the collision avoidance maneuver analysis.

5.2.4 Examples

To demonstrate the explicit expression of P_c , this section will take the U.S. and Russian satellites' collision event for example. Table 5.1 shows the position and velocity coordinates in the J2000 Earth Centered Inertial (ECI) coordinate system at the TCA and eccentricities of two satellites. Both objects are in near-circular orbits.

We now calculate the miss distances, components of relative position in the primary object's RSW system, the altitude difference and time difference, the angle between orbital planes, flight-path angles. The results are shown in Table 5.2.

Chapter 3 has presented a methodology for periodicity characterization and Poisson series fitting for orbital prediction error based on historical orbital data. The Poisson coefficient matrices of each error components are fitted using least squares method. Substituting time difference and mean anomaly into the Poisson series one can obtain the error information at specific time. By using this orbital prediction error analysis method, we obtain the positional error standard deviations of two objects in radial, in-track, and cross-track direction. The combined error standard deviations in each directions and the combined error standard deviation in the horizontal plane are calculated using Eqs. (5.22) and (5.23). The results are shown in Table 5.3.

The aspect ratio of the combined error ellipsoid is $\sigma_R:\sigma_S:\sigma_W = 1:10:1.8$, Fig. 5.11 illustrates the 3σ combined error ellipsoid of two objects at TCA, and Fig. 5.12 shows the combined error ellipse and combined collision cross-section on the conjunction plane.

The collision cross-section radii of two objects are all equal to 5 m, the combined collision cross-section radii are $r_A = 10$ m. In the condition represented by Tables 5.1, 5.2 and 5.3, we calculate P_c of the conjunction cases by means of the general method in Sect. 5.1 and the two explicit expressions in Sect. 5.2. The results are shown in Table 5.4.

The calculated collision probabilities exceed the widely accepted red threshold (10^{-4}) for collision avoidance maneuver.

Table 5.1 Position and velocity coordinates in the ECI system and eccentricities

Object	X/km	Y/km	Z/km	V_x (km/s)	V_y (km/s)	V_z (km/s)	e
Primary	-1457.273246	1589.568484	6814.189959	-7.001731	-2.439512	-0.926209	0.001174
Secondary	-1457.532155	1588.932671	6814.316188	3.578705	-6.172896	2.200215	0.000660

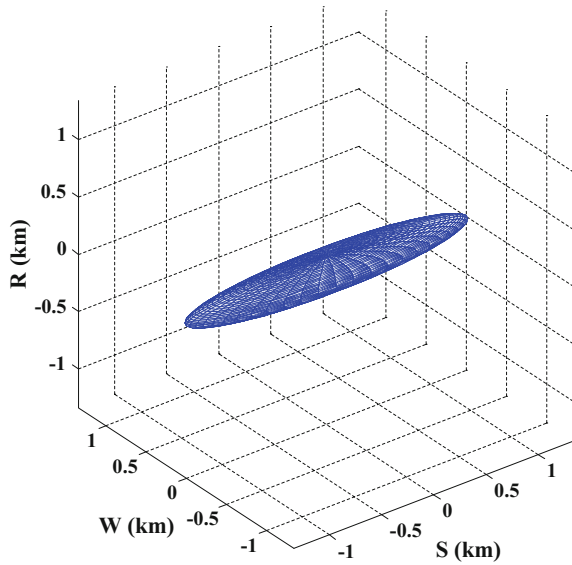
Table 5.2 Conjunction geometry of the U.S. and Russian satellite collision event

Conjunction geometry	Value
Miss distance/km	0.698011
Δh /km	0.031765
Δt /km	0.149075
R /km	0.031731
S /km	0.436476
W /km	0.543785
$\sqrt{S^2 + W^2}$ /km	0.697294
ϕ ($^\circ$)	102.458
θ_1 ($^\circ$)	0.0153
θ_2 ($^\circ$)	-0.0338

Table 5.3 Positional error standard deviations in radial, in-track, cross-track direction and the combined error

Object	σ_R /km	σ_S /km	σ_W /km
Primary	0.0231207	0.2061885	0.0719775
Secondary	0.0363234	0.4102069	0.0341134
Combined	0.0430576	0.4591115	0.0796523
Radial/horizontal	0.0430576	0.2941297	

Fig. 5.11 3σ combined error ellipsoid



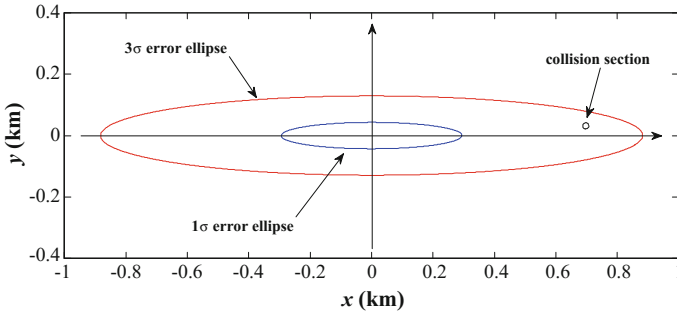


Fig. 5.12 The combined error ellipse and combined collision cross-section on the conjunction plane

Table 5.4 Collision probabilities from Chan’s method and explicit expressions

General method	Explicit expression (conjunction geometries)	Explicit expression (RSW components)
1.814826×10^{-4}	1.806946×10^{-4}	1.807975×10^{-4}

For conjunction in circular orbits, Table 5.4 shows numerical precision to 3 significant figures of both the explicit expressions relative to general method, the relative errors are respectively 0.4342 and 0.3775%. This result reveals that the precisions of both explicit expressions in circular orbits are sufficient.

In addition, using the conjunction geometries described in Table 5.2, we calculate P_c as a function of the size of the covariance. Consider the case in which the covariances of two objects are multiplied by a scaling factor k^2

$$\mathbf{P}_{1RSW}(k) = k^2 \begin{bmatrix} \sigma_{1R}^2 & 0 & 0 \\ 0 & \sigma_{1S}^2 & 0 \\ 0 & 0 & \sigma_{1W}^2 \end{bmatrix}, \quad \mathbf{P}_{2RSW}(k) = k^2 \begin{bmatrix} \sigma_{2R}^2 & 0 & 0 \\ 0 & \sigma_{2S}^2 & 0 \\ 0 & 0 & \sigma_{2W}^2 \end{bmatrix} \quad (5.29)$$

where $\sigma_{1R} \sim \sigma_{2W}$ are given in Table 5.3, and k could be any positive real number. The scaling factor k is chosen as the variable factor. Figure 5.13 depict the P_c curves resulted from three methods with the scaling factor as a parameter for the conjunction cases.

Figure 5.13 shows that in the case of near-circular orbit, the P_c resulted from the conjunction geometries-based explicit expression (denoted by P_{c1} , line with marker ‘ Δ ’) and the RSW components-based explicit expression (denoted by P_{c2} , line with marker ‘ \square ’) match well with the result from the general method (denoted by P_{c0} , line with marker ‘ \circ ’), these lines maybe illegible because they are very close.

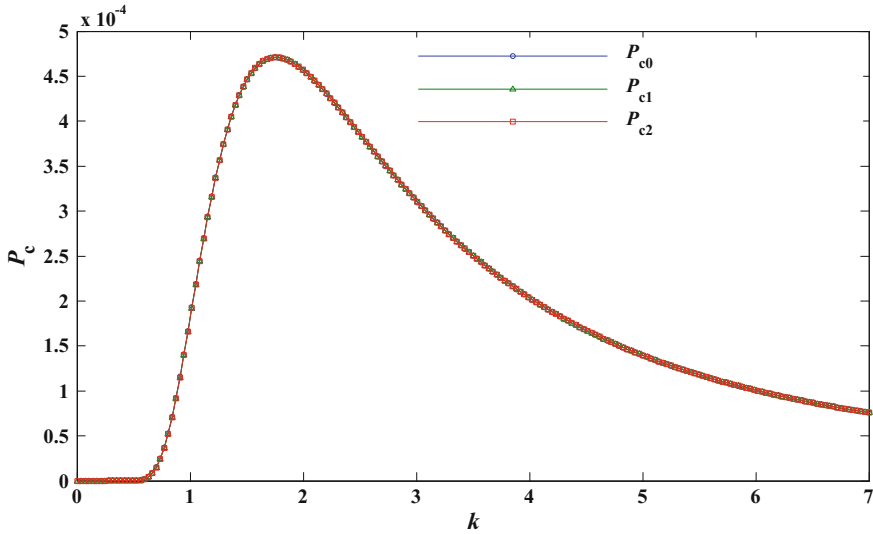


Fig. 5.13 P_c versus error scaling factor using three methods

5.3 Explicit Expression of Collision Probability in General Orbit

The explicit expression of P_c in terms of the relative positions in a circular orbit presented in Sect. 5.2 is simple and compact, reveals the relationship between the collision probability and conjunction geometry more directly. But the circular orbit assumption restricts its application. Besides most orbiting objects in the circular or near-circular orbits, such as Low Earth Orbit (LEO), Middle Earth Orbit (MEO), and Geostationary Earth Orbit (GEO), there are sufficient amount of objects on highly eccentric orbits (HEO). The HEO objects are of special meanings in the manned spacecraft collision avoidance. Manned spacecraft such as Space Shuttle and International Space Station are operated in lower (about 300–500 km) orbits, the lifetime of uncontrolled circular orbit at this region is very short (lifetime of circular orbit at 340 km altitude is about one month). Thus, the population of space objects in circular orbit at this altitude is few, most objects presenting threat to manned spacecrafts are HEO objects with low perigee altitudes. The explicit expression of P_c is not valid for conjunction events in such orbits. In this case, the error of explicit expression of P_c based on the circular orbit assumption is intolerable. Therefore, we need to extend explicit expression of P_c to general orbit type. It is necessary to modify the expression when the eccentricity is not exactly equal to zero.

In the case of general orbit, P_c is expressed as explicit functions of the conjunction geometries (crossing altitude difference and time difference of the common perpendicular line to two velocities, the angle between two velocities, etc.) or the

NTW components of relative position. The explicit expression relates P_c with components of relative position.

5.3.1 Explicit Expression of P_c in Terms of Conjunction Geometries

1. Analysis of closest approach in general orbit

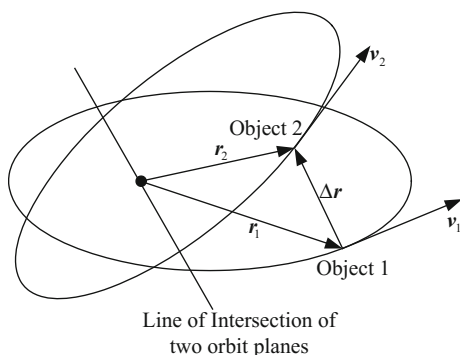
In this section, we will analyze the closest approach between orbits and between objects in general situation. Based on the analysis of closest approach we will give the details of explicit expression of P_c .

Obviously, the collision between space objects always occurs in the vicinity of points of closest approach between their orbits. There will be two closest approach distances between two arbitrary elliptical orbits, as illustrated in Fig. 5.14. It should be noted that the points are not necessarily along the line of intersection of the two planes [49]. In the case of circular orbits, the points of closest approach are along the line of intersection of orbital planes. For elliptical orbits, on the other hand, the points are not along the line of intersection but in the vicinity.

(1) Closest approach between orbits

As two space objects (Object 1 and Object 2) move around their respective orbits, their position vectors and velocity vectors are shown in Fig. 5.14. The closest approach distance between the two elliptical orbits is in nature the closest approach distance between two objects without considering the synchronization of two objects. That means the independent time variables in the motion equation of two objects can be different in the analysis of the orbits' closest approach. In other words, two objects' positions on their respective orbits are arbitrary without any restriction. In this case, closest approach distance between objects would be the closest approach distance between orbits.

Fig. 5.14 Collision between space objects always occurs in the vicinity of line of intersection of their orbit planes



Let state vectors of two objects be $[\mathbf{r}_1(t_1), \mathbf{v}_1(t_1)]$ and $[\mathbf{r}_2(t_2), \mathbf{v}_2(t_2)]$ with different independent variables t_1 and t_2 , relative position vector is given by

$$\Delta \mathbf{r}(t_1, t_2) = \mathbf{r}_2(t_2) - \mathbf{r}_1(t_1) \tag{5.30}$$

The relative separation distance between two objects is the magnitude of relative position vector, the square of relative separation distance can be expressed as

$$\Delta r^2 = \Delta \mathbf{r}(t_1, t_2) \cdot \Delta \mathbf{r}(t_1, t_2) = \mathbf{r}_2(t_2) \cdot \mathbf{r}_2(t_2) - 2\mathbf{r}_1(t_1) \cdot \mathbf{r}_2(t_2) + \mathbf{r}_1(t_1) \cdot \mathbf{r}_1(t_1) \tag{5.31}$$

The partial derivatives of Δr^2 with respect to t_1 and t_2 are

$$\begin{cases} \frac{\partial \Delta r^2}{\partial t_1} = 2\mathbf{r}_1(t_1) \cdot \mathbf{v}_1(t_1) - 2\mathbf{r}_2(t_2) \cdot \mathbf{v}_1(t_1) \\ \frac{\partial \Delta r^2}{\partial t_2} = 2\mathbf{r}_2(t_2) \cdot \mathbf{v}_2(t_2) - 2\mathbf{r}_1(t_1) \cdot \mathbf{v}_2(t_2) \end{cases} \tag{5.32}$$

Let partial derivatives be zeros, the relative minimum distance will be the solution to the two simultaneous equations

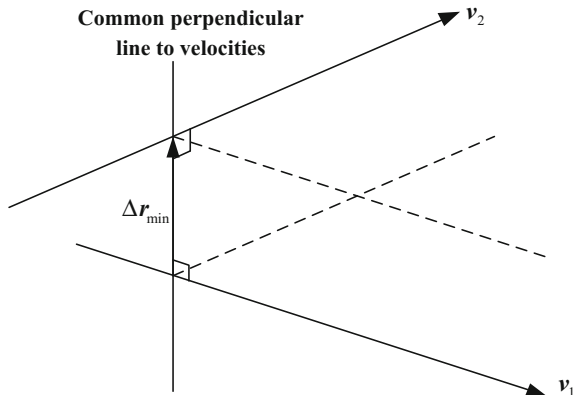
$$\begin{cases} [\mathbf{r}_1(t_1) - \mathbf{r}_2(t_2)] \cdot \mathbf{v}_1(t_1) = 0 \\ [\mathbf{r}_1(t_1) - \mathbf{r}_2(t_2)] \cdot \mathbf{v}_2(t_2) = 0 \end{cases} \tag{5.33}$$

Thus, the minimum relative position vector satisfy

$$\begin{cases} \Delta \mathbf{r}_{\min}(t_1, t_2) \cdot \mathbf{v}_1(t_1) = 0 \\ \Delta \mathbf{r}_{\min}(t_1, t_2) \cdot \mathbf{v}_2(t_2) = 0 \end{cases} \tag{5.34}$$

Equation (5.34) indicates that the necessary condition of closest approach between two orbits is that the relative position vector is perpendicular to both objects' velocity vectors, see Fig. 5.15. The closest approach distance Δr_{\min}

Fig. 5.15 The necessary condition of closest approach between two orbits is that the relative position vector is perpendicular to both objects' velocity vectors



between two orbits is the minimum possible distance between two objects, the actual closest approach distance Δr_{tca} is always smaller than Δr_{min} .

(2) Closest approach between objects

To determine the closest approach distance between two objects at the TCA, the actual positions in their respective orbits must be considered. In this case, there is only one independent time variable. Relative position vector can be written as

$$\Delta \mathbf{r}(t) = \mathbf{r}_2(t) - \mathbf{r}_1(t) \quad (5.35)$$

The square of relative separation distance is

$$\Delta r^2 = \Delta \mathbf{r}(t) \cdot \Delta \mathbf{r}(t) = \mathbf{r}_2(t) \cdot \mathbf{r}_2(t) - 2\mathbf{r}_1(t) \cdot \mathbf{r}_2(t) + \mathbf{r}_1(t) \cdot \mathbf{r}_1(t) \quad (5.36)$$

The derivative of Δr^2 with respect to t is

$$\frac{d\Delta r^2}{dt} = 2(\mathbf{r}_1 \cdot \mathbf{v}_1 + \mathbf{r}_2 \cdot \mathbf{v}_2 - \mathbf{r}_1 \cdot \mathbf{v}_2 - \mathbf{r}_2 \cdot \mathbf{v}_1) \quad (5.37)$$

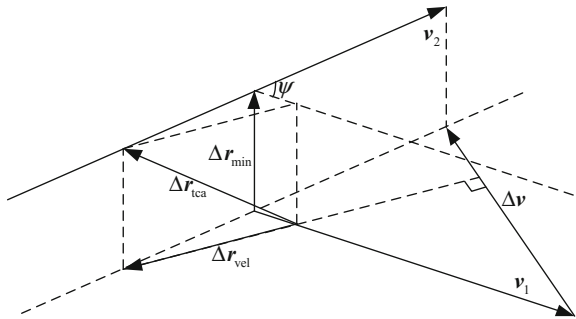
Let the derivative $d\Delta r^2/dt$ be zero, we can obtain

$$(\mathbf{r}_2 - \mathbf{r}_1) \cdot (\mathbf{v}_2 - \mathbf{v}_1) = 0, \quad \Delta \mathbf{r} \cdot \Delta \mathbf{v} = 0 \quad (5.38)$$

Equation (5.38) indicates that the relative position vector of two objects is perpendicular to their relative velocity vector at the TCA. Figure 5.16 illustrates the closest approach between orbits and the closest approach between objects.

Therefore, as shown in Fig. 5.16, relative position vector at the TCA, $\Delta \mathbf{r}_{\text{tca}}$, can be resolved into two components: (1) closest approach vector between orbits, $\Delta \mathbf{r}_{\text{min}}$, which is in the direction of common perpendicular line to two objects' velocity vectors; and (2) closest approach vector in the plane perpendicular to the common perpendicular line to two velocities (referred to as the velocity plane hereafter), $\Delta \mathbf{r}_{\text{vel}}$.

Fig. 5.16 The closest approach vector $\Delta \mathbf{r}_{\text{min}}$ between orbits and the closest approach vector at TCA $\Delta \mathbf{r}_{\text{tca}}$



$$\Delta \mathbf{r}_{\text{tca}} = \Delta \mathbf{r}_{\text{min}} + \Delta \mathbf{r}_{\text{vel}} \quad (5.39)$$

where $\Delta \mathbf{r}_{\text{min}}$ and $\Delta \mathbf{r}_{\text{vel}}$ are perpendicular to each other, so relative distance at the TCA can be given by

$$\Delta r_{\text{tca}}^2 = \Delta r_{\text{min}}^2 + \Delta r_{\text{vel}}^2 \quad (5.40)$$

The analysis states that the relative position vector at the TCA can be resolved to the closest approach distance between two orbits and the time difference crossing the common perpendicular line to two velocities, i.e. the altitude difference and time difference crossing the common perpendicular line to two velocities. This conclusion is valid for any orbit. The similar result is obtained under the circular orbit assumption in Sect. 5.2.1, it is the deduction of this section's conclusion in the case of circular orbit.

(3) Expression of closest approach distance for linear relative motion

In previous section, necessary conditions of closest approach between two elliptical orbits and between objects were discussed in general situation. In short-term conjunction where relative velocity of two objects is sufficiently large, the linear relative motion assumption is always taken when computing the probability of collision. In this section, expressions of closest approach distance between orbits and between objects will be provided by using the position and velocity vectors based on the assumption of linear motion.

As discussed before, the distance between two orbits' paths is minimized when the relative position vector is perpendicular to both objects' velocity vectors, as shown in Fig. 5.15. Under the assumption of linear motion, in the vicinity of the point of closest approach, two objects move along two straight lines and their relative motion trajectory can also be represented by a straight line. The points of closest approach of two straight lines are points of intersection between two lines and their common perpendicular line. We want to know the minimum relative distance between two lines Δr_{min} , and distances from two objects' positions to points of intersection, Δl_1 and Δl_2 . In this way we can obtain the similar parameters as crossing altitude difference and time difference of the line of intersection presented in Sect. 5.2.

Let \mathbf{r}_{10} and \mathbf{r}_{20} are initial position vectors of two objects in the Earth Centered Inertial (ECI) coordinate system. The angle ψ between two velocities is

$$\psi = \arccos \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{|\mathbf{v}_1| |\mathbf{v}_2|} \quad (5.41)$$

The parameterized equations of two linear trajectories are

$$\begin{cases} \mathbf{r}_1(t_1) = \mathbf{r}_{10} + \mathbf{v}_1 t_1 \\ \mathbf{r}_2(t_2) = \mathbf{r}_{20} + \mathbf{v}_2 t_2 \end{cases} \quad (5.42)$$

The relative position vector can be expressed as

$$\Delta \mathbf{r}(t_1, t_2) = \mathbf{r}_2(t_2) - \mathbf{r}_1(t_1) = \mathbf{r}_{20} - \mathbf{r}_{10} + \mathbf{v}_2 t_2 - \mathbf{v}_1 t_1 = \Delta \mathbf{r}_0 + \mathbf{v}_2 t_2 - \mathbf{v}_1 t_1 \quad (5.43)$$

The square of the relative distance is

$$\begin{aligned} \Delta r^2 = & (\Delta \mathbf{r}_0 \cdot \Delta \mathbf{r}_0) + 2(\Delta \mathbf{r}_0 \cdot \mathbf{v}_2)t_2 - 2(\Delta \mathbf{r}_0 \cdot \mathbf{v}_1)t_1 + (\mathbf{v}_2 \cdot \mathbf{v}_2)t_2^2 \\ & - 2(\mathbf{v}_1 \cdot \mathbf{v}_2)t_1 t_2 + (\mathbf{v}_1 \cdot \mathbf{v}_1)t_1^2 \end{aligned} \quad (5.44)$$

Let the derivatives of Δr^2 with respect to t_1 and t_2 equal to zero, then

$$\begin{cases} \frac{\partial \Delta r^2}{\partial t_1} = -(\Delta \mathbf{r}_0 \cdot \mathbf{v}_1) - (\mathbf{v}_1 \cdot \mathbf{v}_2)t_2 + (\mathbf{v}_1 \cdot \mathbf{v}_1)t_1 = 0 \\ \frac{\partial \Delta r^2}{\partial t_2} = (\Delta \mathbf{r}_0 \cdot \mathbf{v}_2) + (\mathbf{v}_2 \cdot \mathbf{v}_2)t_2 - (\mathbf{v}_1 \cdot \mathbf{v}_2)t_1 = 0 \end{cases} \quad (5.45)$$

Then when $t_{1\min}$ and $t_{2\min}$ given by Eq. (5.46) the relative distance will be minimized

$$\begin{cases} t_{1\min} = \frac{(\mathbf{v}_2 \cdot \mathbf{v}_2)(\Delta \mathbf{r}_0 \cdot \mathbf{v}_1) - (\mathbf{v}_1 \cdot \mathbf{v}_2)(\Delta \mathbf{r}_0 \cdot \mathbf{v}_2)}{(\mathbf{v}_1 \cdot \mathbf{v}_1)(\mathbf{v}_2 \cdot \mathbf{v}_2) - (\mathbf{v}_1 \cdot \mathbf{v}_2)(\mathbf{v}_1 \cdot \mathbf{v}_2)} = \frac{\frac{v_2}{v_1}(\Delta \mathbf{r}_0 \cdot \mathbf{v}_1) - \cos \psi (\Delta \mathbf{r}_0 \cdot \mathbf{v}_2)}{v_1 v_2 \sin^2 \psi} \\ t_{2\min} = \frac{(\mathbf{v}_1 \cdot \mathbf{v}_2)(\Delta \mathbf{r}_0 \cdot \mathbf{v}_1) - (\mathbf{v}_1 \cdot \mathbf{v}_1)(\Delta \mathbf{r}_0 \cdot \mathbf{v}_2)}{(\mathbf{v}_1 \cdot \mathbf{v}_1)(\mathbf{v}_2 \cdot \mathbf{v}_2) - (\mathbf{v}_1 \cdot \mathbf{v}_2)(\mathbf{v}_1 \cdot \mathbf{v}_2)} = \frac{\cos \psi (\Delta \mathbf{r}_0 \cdot \mathbf{v}_1) - \frac{v_1}{v_2} (\Delta \mathbf{r}_0 \cdot \mathbf{v}_2)}{v_1 v_2 \sin^2 \psi} \end{cases} \quad (5.46)$$

Let the dimensionless parameter η be the ratio of magnitudes of two velocities

$$\eta \equiv \frac{v_2}{v_1} \quad (5.47)$$

Equation (5.46) could be rewritten as

$$\begin{cases} t_{1\min} = \frac{\eta(\Delta \mathbf{r}_0 \cdot \mathbf{v}_1) - \cos \psi (\Delta \mathbf{r}_0 \cdot \mathbf{v}_2)}{v_1 v_2 \sin^2 \psi} \\ t_{2\min} = \frac{\cos \psi (\Delta \mathbf{r}_0 \cdot \mathbf{v}_1) - (\Delta \mathbf{r}_0 \cdot \mathbf{v}_2)/\eta}{v_1 v_2 \sin^2 \psi} \end{cases} \quad (5.48)$$

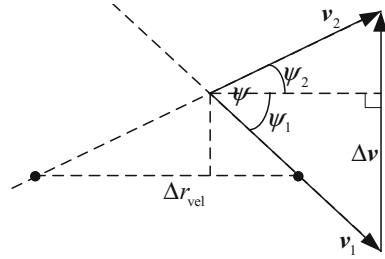
The relative position vector at closest approach between orbits can be expressed as

$$\Delta \mathbf{r}_{\min} = \Delta \mathbf{r}_0 + \mathbf{v}_2 t_{2\min} - \mathbf{v}_1 t_{1\min} \quad (5.49)$$

The closest approach distance between two straight line trajectories will be $\Delta r_{\min} = |\Delta \mathbf{r}_{\min}|$.

From Eq. (5.48) we know that the required moving times from initial positions to points of closest approach between two orbits are $t_{1\min}$ and $t_{2\min}$. Therefore, when

Fig. 5.17 Closest approach distance in the velocity plane



object 1 is arriving at point of intersection, object 2 still need Δt_2 to arrive at the other point of intersection. The time-to-go of object 2 is $\Delta t_2 = t_{2\min} - t_{1\min}$, and corresponding distance between object 2 to point of intersection (distance-to-go) is $\Delta l_2 = v_2 \cdot \Delta t_2$. Similarly, we can obtain time-to-go and distance-to-go of object 1. Let Δt be the crossing time difference of the common perpendicular line

$$\Delta t = \frac{|(\cos \psi - \eta)(\Delta r_0 \cdot v_1) + (\cos \psi - 1/\eta)(\Delta r_0 \cdot v_2)|}{v_1 v_2 \sin^2 \psi} \tag{5.50}$$

As seen in Fig. 5.17, closest approach distance in the velocity plane is determined by time difference of two objects crossing common perpendicular line to velocity vectors, velocity’s magnitude v_1 and v_2 , and the angle between two velocity vectors. We define two angles ψ_1 and ψ_2 as illustrated in Fig. 5.17, the closest approach distance in the velocity plane Δr_{vel} can be expressed as

$$\Delta r_{\text{vel}} = \Delta l_1 \cos \psi_1 = \Delta l_2 \cos \psi_2 \tag{5.51}$$

Using the geometric laws, we know that ψ_1 and ψ_2 satisfy the following formulae with the ratio of magnitude of velocity η and the angle between velocities ψ .

$$\begin{cases} \tan \psi_1 = \frac{1-\eta \cos \psi}{\eta \sin \psi}, \cos \psi_1 = \frac{\eta \sin \psi}{\sqrt{1+\eta^2-2\eta \cos \psi}}, \sin \psi_1 = \frac{1-\eta \cos \psi}{\sqrt{1+\eta^2-2\eta \cos \psi}} \\ \tan \psi_2 = \frac{\eta-\cos \psi}{\sin \psi}, \cos \psi_2 = \frac{\sin \psi}{\sqrt{1+\eta^2-2\eta \cos \psi}}, \sin \psi_2 = \frac{\eta-\cos \psi}{\sqrt{1+\eta^2-2\eta \cos \psi}} \end{cases} \tag{5.52}$$

Thus, Eq. (5.51) can be written as

$$\Delta r_{\text{vel}} = \frac{v_2 \sin \psi \Delta t}{\sqrt{1+\eta^2-2\eta \cos \psi}} \tag{5.53}$$

The closest approach distance between two objects at the TCA is

$$\Delta r_{\text{tca}} = \sqrt{\Delta r_{\text{min}}^2 + \Delta r_{\text{vel}}^2} \tag{5.54}$$

Up to now, we have formulated the closest approach distance at the TCA Δr_{tca} , and its two components Δr_{min} and Δr_{vel} .

The necessary condition of closest approach between two general orbits is that the relative position vector is perpendicular to velocity vectors of both objects, i.e. the relative position vector is parallel to the common perpendicular line to two velocities. The necessary condition of closest approach between two space objects moving around their orbits is that the relative position vector is perpendicular to relative velocity vector. Based on these conclusions the relative position vector at the time of closest approach (TCA) can be resolved to two components: (1) closest approach vector between orbits that is in the direction of common perpendicular line to velocity vectors of both objects, and (2) closest approach vector in the plane perpendicular to the common perpendicular line to two velocities. The conjunction geometry of two objects can be represented by altitude difference and time difference crossing common perpendicular line to velocities. Under the linear motion assumption, closest approach distance between orbits and between objects are formulated using position and velocity coordinates of two objects at conjunction.

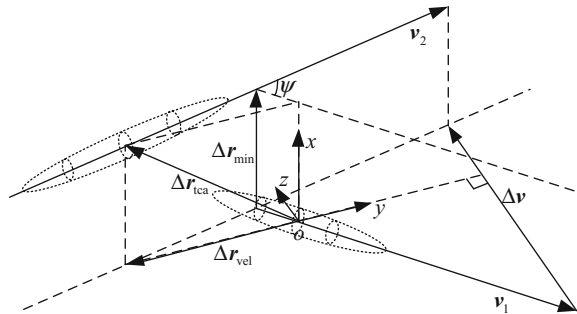
2. Definition of encounter coordinate and projection of position error covariance

Based on the analysis of closest approach distance, we can define the encounter coordinate system, $o-xyz$. The origin o is at the centroid of object 1, the z -axis points in the direction of relative velocity Δv , the x -axis points in the direction of closest approach vector between orbits Δr_{min} , the y -axis lies in the velocity plane and perpendicular to the relative velocity vector. The encounter coordinate system has a right-handed sense. The x - y plane perpendicular to the relative velocity is the conjunction plane, for the closest approach vector Δr_{tca} is in this plane, as shown in Fig. 5.18.

At the TCA, the coordinates of distributing center of the position error ellipsoid of object 2 in the encounter coordinate system are

$$\begin{cases} \mu_x = \Delta r_{\text{min}} \\ \mu_y = \Delta r_{\text{vel}} = \frac{v_2 \sin \psi \Delta t}{\sqrt{1 + \eta^2 - 2\eta \cos \psi}} \end{cases} \quad (5.55)$$

Fig. 5.18 Definition of the encounter coordinate system and position error ellipsoids of two objects



For objects in circular orbit, the position error covariance is described in the RSW satellite-based coordinate system. For objects in the HEO orbit, it is more advantageous to describe position errors in the NTW satellite-based coordinate system, which is valid for all orbit types, than in the RSW coordinate system. In the NTW coordinate system, the T axis (tangential or in-track) is tangential to the orbit and always points to the velocity vector. The N axis (normal) lies in the orbital plane, normal to the velocity vector. The W axis (cross-track) is normal to the orbital plane. The NTW and the RSW systems are identical for circular orbits. In order to project position error covariance to the encounter coordinate system, we must know the transformations from the NTW coordinate to the encounter coordinate. Obviously, the N axis is not parallel to vector $\Delta \mathbf{r}_{\min}$, we assume that the N axis points to $\Delta \mathbf{r}_{\min}$, hence we just need to rotate the NTW systems of two objects about their N axis to get the encounter coordinate system. The angles of rotation are ψ_1 and $-\psi_2$ respectively (see Fig. 5.18), the coordinate transformation matrices are given in Eq. (5.56).

$$\begin{cases} \mathbf{M}_{1\text{NTW} \rightarrow \text{xyz}} = \mathbf{M}_1(\psi_1) \\ \mathbf{M}_{2\text{NTW} \rightarrow \text{xyz}} = \mathbf{M}_1(-\psi_2) \end{cases} \quad (5.56)$$

The position error covariance matrices associated with the objects are described in their satellite-based NTW coordinate system. It is reasonable to assume that the covariance matrices are both diagonal (see Fig. 5.18), given by

$$\begin{cases} \mathbf{P}_{1\text{NTW}} = \begin{bmatrix} \sigma_{1\text{N}}^2 & 0 & 0 \\ 0 & \sigma_{1\text{T}}^2 & 0 \\ 0 & 0 & \sigma_{1\text{W}}^2 \end{bmatrix} \\ \mathbf{P}_{2\text{NTW}} = \begin{bmatrix} \sigma_{2\text{N}}^2 & 0 & 0 \\ 0 & \sigma_{2\text{T}}^2 & 0 \\ 0 & 0 & \sigma_{2\text{W}}^2 \end{bmatrix} \end{cases} \quad (5.57)$$

Transforming $\mathbf{P}_{1\text{NTW}}$ and $\mathbf{P}_{2\text{NTW}}$ to the encounter coordinate system by using transformation matrices in Eq. (5.56), we can obtain

$$\begin{aligned} \mathbf{P}_1 &= \mathbf{M}_{1\text{NTW} \rightarrow \text{xyz}} \mathbf{P}_{1\text{NTW}} \mathbf{M}_{1\text{NTW} \rightarrow \text{xyz}}^T \\ &= \begin{bmatrix} \sigma_{1\text{N}}^2 & 0 & 0 \\ 0 & \sigma_{1\text{T}}^2 \cos^2 \psi_1 + \sigma_{1\text{W}}^2 \sin^2 \psi_1 & (\sigma_{1\text{W}}^2 - \sigma_{1\text{T}}^2) \sin \psi_1 \cos \psi_1 \\ 0 & (\sigma_{1\text{W}}^2 - \sigma_{1\text{T}}^2) \sin \psi_1 \cos \psi_1 & \sigma_{1\text{T}}^2 \sin^2 \psi_1 + \sigma_{1\text{W}}^2 \cos^2 \psi_1 \end{bmatrix} \end{aligned} \quad (5.58)$$

$$\begin{aligned} \mathbf{P}_2 &= \mathbf{M}_{2\text{NTW} \rightarrow \text{xyz}} \mathbf{P}_{2\text{NTW}} \mathbf{M}_{2\text{NTW} \rightarrow \text{xyz}}^T \\ &= \begin{bmatrix} \sigma_{2\text{N}}^2 & 0 & 0 \\ 0 & \sigma_{2\text{T}}^2 \cos^2 \psi_2 + \sigma_{2\text{W}}^2 \sin^2 \psi_2 & (\sigma_{2\text{T}}^2 - \sigma_{2\text{W}}^2) \sin \psi_2 \cos \psi_2 \\ 0 & (\sigma_{2\text{T}}^2 - \sigma_{2\text{W}}^2) \sin \psi_2 \cos \psi_2 & \sigma_{2\text{T}}^2 \sin^2 \psi_2 + \sigma_{2\text{W}}^2 \cos^2 \psi_2 \end{bmatrix} \end{aligned} \quad (5.59)$$

We just take two components in the encounter plane into account, and sum-up two covariance matrices to get joint error covariance matrix associated with the relative position between two objects

$$\mathbf{P} = \begin{bmatrix} \sigma_{1N}^2 + \sigma_{2N}^2 & 0 \\ 0 & \sigma_{1T}^2 \cos^2 \psi_1 + \sigma_{1W}^2 \sin^2 \psi_1 + \sigma_{2T}^2 \cos^2 \psi_2 + \sigma_{2W}^2 \sin^2 \psi_2 \end{bmatrix} \quad (5.60)$$

Combining Eqs. (5.52) and (5.60), the 2-dimensional joint error covariance matrix in encounter plane can be written as

$$\mathbf{P} = \begin{bmatrix} \sigma_{1N}^2 + \sigma_{2N}^2 & 0 \\ 0 & \frac{\sigma_{1T}^2 \eta^2 \sin^2 \psi + (1 - \eta \cos \psi)^2 \sigma_{1W}^2 + \sigma_{2T}^2 \sin^2 \psi + (\eta - \cos \psi)^2 \sigma_{2W}^2}{1 + \eta^2 - 2\eta \cos \psi} \end{bmatrix} \quad (5.61)$$

The joint error covariance matrix \mathbf{P} in the encounter plane is also diagonal, so the P_c density function can be integrated in the encounter plane. The joint error variances in the x and y directions in the encounter plane are

$$\begin{cases} \sigma_x^2 = \sigma_{1N}^2 + \sigma_{2N}^2 \\ \sigma_y^2 = \frac{\sigma_{1T}^2 \eta^2 \sin^2 \psi + (1 - \eta \cos \psi)^2 \sigma_{1W}^2 + \sigma_{2T}^2 \sin^2 \psi + (\eta - \cos \psi)^2 \sigma_{2W}^2}{1 + \eta^2 - 2\eta \cos \psi} \end{cases} \quad (5.62)$$

3. Explicit expression of collision probability

Substituting Eqs. (5.55) and (5.62) into Eq. (5.9), we can obtain the explicit expression of P_c

$$P_c = \exp \left[-\frac{1}{2} \left(\frac{\mu_x^2}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2} \right) \right] \left[1 - \exp \left(-\frac{r_A^2}{2\sigma_x \sigma_y} \right) \right]$$

$$\mu_x = \Delta r_{\min}$$

$$\mu_y = \frac{v_2 \sin \psi \Delta t}{\sqrt{1 + \eta^2 - 2\eta \cos \psi}} \quad (5.63)$$

$$\sigma_x^2 = \sigma_{1N}^2 + \sigma_{2N}^2$$

$$\sigma_y^2 = \frac{\sigma_{1T}^2 \eta^2 \sin^2 \psi + (1 - \eta \cos \psi)^2 \sigma_{1W}^2 + \sigma_{2T}^2 \sin^2 \psi + (\eta - \cos \psi)^2 \sigma_{2W}^2}{1 + \eta^2 - 2\eta \cos \psi}$$

where Δr_{\min} is the closest approach distance between orbits; Δt is the crossing time difference of the common perpendicular line to two objects' velocities; ψ is the angle between two velocity vectors; v_1 and v_2 are the velocities of two objects; η is the ratio of the magnitudes of two velocities; $(\sigma_{1N}^2, \sigma_{1T}^2, \sigma_{1W}^2)$ and $(\sigma_{2N}^2, \sigma_{2T}^2, \sigma_{2W}^2)$ are position error standard deviations of two objects in their NTW coordinate system; r_A is the sum of two objects' equivalent radii. Equation (5.63) expresses the

P_c explicitly with encounter geometries and position error in their NTW coordinate systems.

Furthermore, Eq. (5.63) indicates that the position error in N direction is not coupled with position error in T and W direction; one can sum up error variances of two objects in N direction to obtain the joint error variance

$$\sigma_N^2 = \sigma_{1N}^2 + \sigma_{2N}^2 \quad (5.64)$$

Whereas position errors in T and W direction of two objects are coupled by their velocities and relative velocity, we define joint error variance in the velocity plane as

$$\sigma_V^2 = \frac{\sigma_{1T}^2 \eta^2 \sin^2 \psi + (1 - \eta \cos \psi)^2 \sigma_{1W}^2 + \sigma_{2T}^2 \sin^2 \psi + (\eta - \cos \psi)^2 \sigma_{2W}^2}{1 + \eta^2 - 2\eta \cos \psi} \quad (5.65)$$

From Eq. (5.63) we know that P_c can be considered as a function of two error variances, σ_N^2 and σ_V^2 , which are mutually independent. The two error variances are related to closest approach distance between orbits, Δr_{\min} , and crossing time difference of common perpendicular line to two velocity, Δt , respectively.

Based on the analysis of closest approach, the encounter coordinate system is defined and the transformations from the NTW coordinate systems of two objects to the encounter coordinate system are provided. The position error covariance associated with two objects described in the NTW coordinate system are projected to the encounter coordinate, the explicit expression of P_c is deduced. The P_c is expressed as an explicit function of the encounter geometry (crossing altitude difference and time difference of the common perpendicular line to velocities, the angle between two velocities, magnitude of the relative velocity, etc.) and position error variances in the NTW system.

During the derivation of explicit expression, the assumption that the N axis and $\Delta \mathbf{r}_{\min}$ is parallel was made. This assumption will introduce error. The explicit expression of P_c in terms of the NTW components of relative position can solve this problem well.

5.3.2 *Explicit Expression of P_c in Terms of NTW Components*

The explicit expression of P_c in terms of the relative positions in a circular orbit presented in Sect. 5.2 is simple and compact, reveals the relationship between the collision probability and conjunction geometry more directly. But the circular orbit assumption restricts its application. Although most space objects, especially in LEO and GEO regions, are in circular or near-circular orbits, it is still necessary to

modify the expression when the eccentricity is not exactly equal to zero. In this section, we will give the modified expression of P_c in an eccentric orbit.

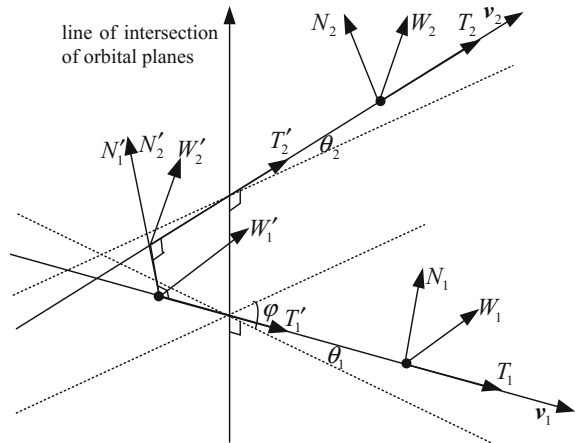
During the derivation of explicit expression of P_c the circular orbit assumption have two major inferences: (1) the velocity vectors of two objects are perpendicular to their position vectors, i.e. the flight-path angles are zeros; (2) the velocities' magnitudes of two objects are equal. The non-zero flight-path angle and unequal velocity's magnitude must be introduced to modify the explicit expression to be suitable for eccentric orbit.

1. Definition of coordinate systems and their transformations

The conjunction geometry near the line of intersection of orbital planes of two objects in general orbit is illustrated in Fig. 5.19. The flight-path angles of two objects are θ_1 and θ_2 , the magnitudes of two velocities are v_1 and v_2 , respectively. The angle between two orbital planes is φ . Five coordinate systems could be defined as illustrated in Fig. 5.19. These five systems are fixed in the inertial space and do not move with the objects.

- (1) Primary NTW system $o_1-N_1T_1W_1$: The origin is at the primary objects at TCA, the T_1 -axis is tangential to the orbit and always points to the primary velocity vector. The N_1 -axis lies in the orbital plane, normal to the velocity vector. The W_1 -axis is normal to the primary orbital plane.
- (2) Secondary NTW system $o_2-N_2T_2W_2$: The origin is at the secondary objects at TCA, the T_2 -axis is tangential to the orbit and always points to the secondary velocity vector. The N_2 -axis lies in the orbital plane, normal to the velocity vector. The W_2 -axis is normal to the secondary orbital plane.
- (3) Modified primary NTW system $o'_1-N'_1T'_1W'_1$: The origin is at the intersection point of the common perpendicular line to two velocities and primary object's orbit. The N'_1 -axis points in the direction of the common perpendicular line vector, the T'_1 -axis points to the primary velocity vector.

Fig. 5.19 The conjunction geometry and coordinate systems



- (4) Modified secondary NTW system $o'_2-N'_2T'_2W'_2$: The origin is at the intersection point of the common perpendicular line to two velocities and secondary object's orbit. The N'_2 -axis points in the direction of the common perpendicular line vector, same as the N'_1 -axis, the T'_2 -axis points to the secondary velocity vector.
- (5) Encounter system $o-xyz$: The origin is at o'_1 of $o'_1-N'_1T'_1W'_1$ system. The x -axis points to the N'_1 -axis of the $o'_1-N'_1T'_1W'_1$ system, the z -axis points to the relative velocity vector, the x - y plane perpendicular to the relative velocity is the encounter plane. This system is not illustrated in Fig. 5.19.

The T_1 -axis of primary NTW system and T'_1 -axis of modified primary NTW system are identical, so we just need to rotate the NTW system about its T_1 -axis with an angle ϕ_1 for transforming from NTW system to modified NTW system.

$$\mathbf{M}_{N_1T_1W_1 \rightarrow N'_1T'_1W'_1} = \mathbf{M}_2(\phi_1) = \begin{bmatrix} \cos \phi_1 & 0 & -\sin \phi_1 \\ 0 & 1 & 0 \\ \sin \phi_1 & 0 & \cos \phi_1 \end{bmatrix} \quad (5.66)$$

The N'_1 -axis of modified primary NTW system and x -axis of encounter system are identical, so we just need to rotate the modified NTW system about its N'_1 -axis with an angle ψ_1 for transforming from modified NTW system to encounter system.

$$\mathbf{M}_{N'_1T'_1W'_1 \rightarrow xyz} = \mathbf{M}_1(\psi_1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \psi_1 & \sin \psi_1 \\ 0 & -\sin \psi_1 & \cos \psi_1 \end{bmatrix} \quad (5.67)$$

Hence the transformation matrix from primary NTW system to encounter system is

$$\begin{aligned} \mathbf{M}_{N_1T_1W_1 \rightarrow xyz} &= \mathbf{M}_{N'_1T'_1W'_1 \rightarrow xyz} \mathbf{M}_{N_1T_1W_1 \rightarrow N'_1T'_1W'_1} \\ &= \begin{bmatrix} \cos \phi_1 & 0 & -\sin \phi_1 \\ \sin \psi_1 \sin \phi_1 & \cos \psi_1 & \sin \psi_1 \cos \phi_1 \\ \cos \psi_1 \sin \phi_1 & -\sin \psi_1 & \cos \psi_1 \cos \phi_1 \end{bmatrix} \end{aligned} \quad (5.68)$$

Similarly, the transformation matrix from secondary NTW system to encounter system is

$$\begin{aligned} \mathbf{M}_{N_2T_2W_2 \rightarrow xyz} &= \mathbf{M}_{N'_2T'_2W'_2 \rightarrow xyz} \mathbf{M}_{N_2T_2W_2 \rightarrow N'_2T'_2W'_2} \\ &= \begin{bmatrix} \cos \phi_2 & 0 & -\sin \phi_2 \\ \sin \psi_2 \sin \phi_2 & \cos \psi_2 & \sin \psi_2 \cos \phi_2 \\ \cos \psi_2 \sin \phi_2 & -\sin \psi_2 & \cos \psi_2 \cos \phi_2 \end{bmatrix} \end{aligned} \quad (5.69)$$

The transformation from each NTW system to encounter system is governed by two angles: ϕ and ψ . These angles will be discussed later.

2. Relative position vector and covariance in encounter system

Suppose the relative position vectors at the TCA in the primary NTW system is

$$\Delta \mathbf{r}_{N_1 T_1 W_1} = [N \quad T \quad W]^T \quad (5.70)$$

so the relative position vectors in the encounter system will be

$$\Delta \mathbf{r}_{xyz} = \mathbf{M}_{N_1 T_1 W_1 \rightarrow xyz} \Delta \mathbf{r}_{N_1 T_1 W_1} = \begin{bmatrix} N \cos \phi_1 - W \sin \phi_1 \\ N \sin \psi_1 \sin \phi_1 + T \cos \psi_1 + W \sin \psi_1 \cos \phi_1 \\ N \cos \psi_1 \sin \phi_1 - T \sin \psi_1 + W \cos \psi_1 \cos \phi_1 \end{bmatrix} \quad (5.71)$$

The position error covariance matrices associated with the objects in eccentric orbit are generally defined in their object-centered NTW coordinate system. It is reasonable to assume that the covariance matrices are both diagonal, given by Eq. (5.57). Transforming \mathbf{P}_{1NTW} and \mathbf{P}_{2NTW} to the encounter coordinate system by using transformation matrices in Eqs. (5.68) and (5.69), we can obtain

$$\begin{aligned} \mathbf{P}_1 &= \mathbf{M}_{N_1 T_1 W_1 \rightarrow xyz} \mathbf{P}_{1NTW} \mathbf{M}_{N_1 T_1 W_1 \rightarrow xyz}^T \\ &= \begin{bmatrix} \sigma_{1N}^2 \cos^2 \phi_1 + \sigma_{1W}^2 \sin^2 \phi_1 & \frac{\sigma_{1N}^2 - \sigma_{1W}^2}{2} \sin \psi_1 \sin 2\phi_1 & \frac{\sigma_{1N}^2 - \sigma_{1W}^2}{2} \cos \psi_1 \sin 2\phi_1 \\ \frac{\sigma_{1N}^2 - \sigma_{1W}^2}{2} \sin \psi_1 \sin 2\phi_1 & (\sigma_{1N}^2 \sin^2 \phi_1 + \sigma_{1W}^2 \cos^2 \phi_1) \sin^2 \psi_1 + \sigma_{1T}^2 \cos^2 \psi_1 & \frac{\sin 2\psi_1}{2} (\sigma_{1N}^2 \sin^2 \phi_1 - \sigma_{1T}^2 + \sigma_{1W}^2 \cos^2 \phi_1) \\ \frac{\sigma_{1N}^2 - \sigma_{1W}^2}{2} \cos \psi_1 \sin 2\phi_1 & \frac{\sin 2\psi_1}{2} (\sigma_{1N}^2 \sin^2 \phi_1 - \sigma_{1T}^2 + \sigma_{1W}^2 \cos^2 \phi_1) & (\sigma_{1N}^2 \sin^2 \phi_1 + \sigma_{1W}^2 \cos^2 \phi_1) \cos^2 \psi_1 + \sigma_{1T}^2 \sin^2 \psi_1 \end{bmatrix} \\ \mathbf{P}_2 &= \mathbf{M}_{N_2 T_2 W_2 \rightarrow xyz} \mathbf{P}_{2NTW} \mathbf{M}_{N_2 T_2 W_2 \rightarrow xyz}^T \\ &= \begin{bmatrix} \sigma_{2N}^2 \cos^2 \phi_2 + \sigma_{2W}^2 \sin^2 \phi_2 & \frac{\sigma_{2N}^2 - \sigma_{2W}^2}{2} \sin \psi_2 \sin 2\phi_2 & \frac{\sigma_{2N}^2 - \sigma_{2W}^2}{2} \cos \psi_2 \sin 2\phi_2 \\ \frac{\sigma_{2N}^2 - \sigma_{2W}^2}{2} \sin \psi_2 \sin 2\phi_2 & (\sigma_{2N}^2 \sin^2 \phi_2 + \sigma_{2W}^2 \cos^2 \phi_2) \sin^2 \psi_2 + \sigma_{2T}^2 \cos^2 \psi_2 & \frac{\sin 2\psi_2}{2} (\sigma_{2N}^2 \sin^2 \phi_2 - \sigma_{2T}^2 + \sigma_{2W}^2 \cos^2 \phi_2) \\ \frac{\sigma_{2N}^2 - \sigma_{2W}^2}{2} \cos \psi_2 \sin 2\phi_2 & \frac{\sin 2\psi_2}{2} (\sigma_{2N}^2 \sin^2 \phi_2 - \sigma_{2T}^2 + \sigma_{2W}^2 \cos^2 \phi_2) & (\sigma_{2N}^2 \sin^2 \phi_2 + \sigma_{2W}^2 \cos^2 \phi_2) \cos^2 \psi_2 + \sigma_{2T}^2 \sin^2 \psi_2 \end{bmatrix} \end{aligned} \quad (5.72)$$

We just take two components in the encounter plane into account, and sum-up two covariance matrices to get combined error covariance matrix associated with the relative position between two objects, $\mathbf{P} = \mathbf{P}_1 + \mathbf{P}_2$. The four elements of the combined error covariance matrix \mathbf{P} are

$$\begin{cases} P_{11} = \sigma_{1N}^2 \cos^2 \phi_1 + \sigma_{2N}^2 \cos^2 \phi_2 + \sigma_{1W}^2 \sin^2 \phi_1 + \sigma_{2W}^2 \sin^2 \phi_2 \\ P_{12} = P_{21} = \frac{\sigma_{1N}^2 - \sigma_{1W}^2}{2} \sin \psi_1 \sin 2\phi_1 + \frac{\sigma_{2N}^2 - \sigma_{2W}^2}{2} \sin \psi_2 \sin 2\phi_2 \\ P_{22} = (\sigma_{1N}^2 \sin^2 \phi_1 + \sigma_{1W}^2 \cos^2 \phi_1) \sin^2 \psi_1 + \sigma_{1T}^2 \cos^2 \psi_1 \\ \quad + (\sigma_{2N}^2 \sin^2 \phi_2 + \sigma_{2W}^2 \cos^2 \phi_2) \sin^2 \psi_2 + \sigma_{2T}^2 \cos^2 \psi_2 \end{cases} \quad (5.73)$$

In most cases the error variance in the normal and cross-track directions are approximately equal while error in the tangential direction is remarkably larger than other two directions. The normal and cross-track components of covariance are assumed to be equal here, i.e. $\sigma_{1N}^2 \approx \sigma_{1W}^2$ and $\sigma_{2N}^2 \approx \sigma_{2W}^2$. Therefore the off-diagonal elements of Eq. (5.73) are kind of inessential compared to the diagonal

elements. In the following analysis we will ignore them. The results in next section will validate this approximation.

So far, we have obtained analytical expression of probability integral parameters $\mu_x, \mu_y, \sigma_x,$ and σ_y from Eqs. (5.71) and (5.73). We can express the P_c as

$$P_c = \exp \left[-\frac{1}{2} \left(\frac{\mu_x^2}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2} \right) \right] \left[1 - \exp \left(-\frac{r_A^2}{2\sigma_x\sigma_y} \right) \right] \tag{5.74}$$

where

$$\begin{cases} \mu_x = N \cos \phi_1 - W \sin \phi_1 \\ \mu_y = N \sin \psi_1 \sin \phi_1 + T \cos \psi_1 + W \sin \psi_1 \cos \phi_1 \\ \sigma_x^2 = \sigma_{1N}^2 \cos^2 \phi_1 + \sigma_{2N}^2 \cos^2 \phi_2 + \sigma_{1W}^2 \sin^2 \phi_1 + \sigma_{2W}^2 \sin^2 \phi_2 \\ \sigma_y^2 = (\sigma_{1N}^2 \sin^2 \phi_1 + \sigma_{1W}^2 \cos^2 \phi_1) \sin^2 \psi_1 + \sigma_{1T}^2 \cos^2 \psi_1 \\ \quad + (\sigma_{2N}^2 \sin^2 \phi_2 + \sigma_{2W}^2 \cos^2 \phi_2) \sin^2 \psi_2 + \sigma_{2T}^2 \cos^2 \psi_2 \end{cases}$$

Moreover, if one wants to get more accurate result, the probability integral parameters $(\mu_x, \mu_y, \sigma_x, \sigma_y)$ in Eq. (5.74) could be substituted to Eq. (5.7) instead of Eq. (5.74) and the term number m in Eq. (5.7) could be set to 1, 2, or larger numbers.

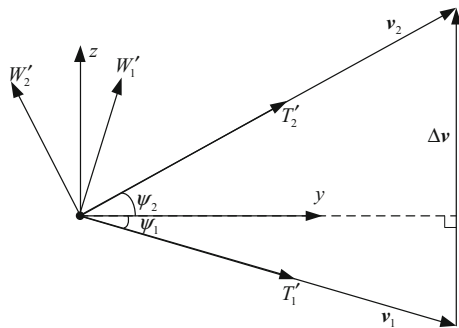
3. Expression of rotating angles

The angle ψ between two velocities can be expressed by using θ_1, θ_2 and φ as

$$\psi = \arccos(\sin \theta_1 \sin \theta_2 + \cos \theta_1 \cos \theta_2 \cos \varphi) \tag{5.75}$$

Figure 5.20 shows the modified primary and secondary NTW systems and the encounter system in the plane perpendicular to the x -axis, where ψ_1 and ψ_2 are the angles from primary and secondary velocity (i.e. the T'_1 -axis and T'_2 -axis) to the y -axis of the encounter system, which can be determined by the angle ψ between two velocities and the ratio of two velocities' magnitudes $\eta = v_2/v_1$ as in Eq. (5.52).

Fig. 5.20 Transformation from modified NTW systems to the encounter system



The rotating angles ϕ_1 and ϕ_2 from the primary and secondary NTW systems to corresponding modified NTW systems can be determined by the angle φ between two orbital planes and the flight-path angles θ_1, θ_2 of two objects.

$$\begin{cases} \tan \phi_1 = \frac{\cos \theta_1 \sin \theta_2 - \sin \theta_1 \cos \theta_2 \cos \varphi}{\cos \theta_2 \sin \varphi} \\ \tan \phi_2 = \frac{\cos \theta_2 \sin \theta_1 - \sin \theta_2 \cos \theta_1 \cos \varphi}{\cos \theta_1 \sin \varphi} \end{cases} \quad (5.76)$$

Summing up Eqs. (5.52), (5.75), (5.76) we obtain the relationship between the rotating angles $(\phi_1, \phi_2, \psi_1, \psi_2)$ and the conjunction geometry $(\varphi, \theta_1, \theta_2, \eta)$.

$$\begin{aligned} \psi &= \arccos(\sin \theta_1 \sin \theta_2 + \cos \theta_1 \cos \theta_2 \cos \varphi) \\ \tan \psi_1 &= \frac{1 - \eta \cos \psi}{\eta \sin \psi} \\ \tan \psi_2 &= \frac{\eta - \cos \psi}{\sin \psi} \\ \tan \phi_1 &= \frac{\cos \theta_1 \sin \theta_2 - \sin \theta_1 \cos \theta_2 \cos \varphi}{\cos \theta_2 \sin \varphi} \\ \tan \phi_2 &= \frac{\cos \theta_2 \sin \theta_1 - \sin \theta_2 \cos \theta_1 \cos \varphi}{\cos \theta_1 \sin \varphi} \end{aligned} \quad (5.77)$$

So far, Eqs. (5.74) and (5.77) express the P_c in general orbit as analytical equation in terms of the relative positions (the N, T, W components) at the TCA, the angle φ between two orbital planes, the flight-path angles θ_1 and θ_2 of two objects, and the ratio η of two velocities' magnitudes.

This modified expression of P_c containing flight-path angle and ratio of velocity's magnitude is subject only to the assumptions that the principal axes of the primary and secondary covariance matrices are aligned with their NTW systems. There are no restrictions imposed on the shapes or eccentricities of objects' orbits, so it is robust and applicable.

Furthermore, if both objects are in circular orbit, i.e. $\theta_1 = \theta_2 = 0$ and $\eta = 1$, the rotating angles would be degraded to

$$\psi = \varphi, \quad \psi_1 = \psi_2 = \frac{\varphi}{2}, \quad \phi_1 = \phi_2 = 0 \quad (5.78)$$

So Eq. (5.74) can be written as

$$P_c = \exp \left[-\frac{1}{2} \left(\frac{\mu_x^2}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2} \right) \right] \left[1 - \exp \left(-\frac{r_A^2}{2\sigma_x \sigma_y} \right) \right] \quad (5.79)$$

where

$$\begin{cases} \mu_x = N \\ \mu_y = T^2 + W^2 \\ \sigma_x^2 = \sigma_{1N}^2 + \sigma_{2N}^2 \\ \sigma_y^2 = (\sigma_{1T}^2 + \sigma_{2T}^2) \cos^2 \frac{\varphi}{2} + (\sigma_{1W}^2 + \sigma_{2W}^2) \sin^2 \frac{\varphi}{2} \end{cases}$$

This equation coincident with Eq. (5.27) is the explicit expression of P_c in circular orbit in terms of the components of relative position.

5.3.3 Examples

On 12 March the crew of the International Space Station (ISS) temporarily retreated into the safety of their Soyuz TMA-13 spacecraft when a small piece of orbital debris (International Designator 1993-032D, U.S. Satellite Number 25090) was belatedly projected to come close to the ISS. At the time of the conjunction, ISS was in a nearly circular orbit near 355 km altitude, while the debris, a piece of mission-related hardware from a U.S. Delta 2 third stage, was in a highly elliptical orbit of about 145 km by 4230 km. From radar cross-section observations, the size of the object was determined to be about 13 cm, large enough to inflict serious damage to the ISS in the event of a collision [50].

To demonstrate the explicit expression of P_c in general orbit, this section will take this event for example. Table 5.5 shows the position and velocity coordinates in the J2000 Earth Centered Inertial (ECI) coordinate system at the TCA and eccentricities of two objects. The primary object is in a near-circular orbit but the secondary object is in a highly eccentric orbit.

We now calculate the miss distances, components of relative position in the primary object’s NTW system, the altitude difference and time difference, the angle between orbital planes, flight-path angles, and ratio of velocity’s magnitude. The results are shown in Table 5.6.

By using the orbital prediction error analysis method in Chap. 3, we obtain the positional error standard deviations of two objects in normal, tangential, and cross-track direction. The combined error standard deviations in each directions and the combined error standard deviation in the horizontal plane are calculated using Eq. (5.74). The results are shown in Table 5.7.

Let the collision cross-section radii of ISS and the debris equal to 100 and 0.13 m, respectively. The combined collision cross-section radii are $r_A =$

Table 5.5 Position and velocity coordinates in the ECI system and eccentricities (ISS event)

Object	X/km	Y/km	Z/km	V_x /(km/s)	V_y /(km/s)	V_z /(km/s)	e
Primary	3126.0188	5227.1461	-2891.3029	-3.2980	4.7587	5.0543	0.001300
Secondary	3124.3685	5226.0042	-2889.9446	-7.7726	1.9308	-2.7580	0.238283

Table 5.6 Conjunction geometry of the ISS event

Conjunction geometry	Value
Miss distance/km	2.423292
$\Delta r_{\min}/\text{km}$	2.182973
$\Delta t/\text{km}$	0.161021
N/km	-2.232155
T/km	0.894487
W/km	-0.299533
$\varphi/(\text{°})$	71.179
$\theta_1/(\text{°})$	-0.0538
$\theta_2/(\text{°})$	-6.258
η	1.102111

Table 5.7 Positional error standard deviations in normal, tangential, cross-track direction and the combined error (ISS event)

Object	σ_N/km	σ_T/km	σ_W/km
Primary	0.5548968	6.185655	1.9433925
Secondary	0.8717616	12.306207	0.9210618
Combined	1.055317	10.9156858	

100.13 m. In the condition represented by Tables 5.5, 5.6 and 5.7, we calculate P_c of the conjunction cases by means of the general method in Sect. 5.1 and the two explicit expressions in Sect. 5.3. The results are shown in Table 5.8.

For conjunction in the eccentric orbits (ISS event), Table 5.8 shows that the relative error of the results from explicit expression in terms of the conjunction geometries and in terms of the NTW components to the result from general method are 6.51 and 0.3433%, respectively. This result reveals that the precisions of both explicit expressions in general orbits are sufficient.

In addition, using the conjunction geometries described in Table 5.6, we calculate P_c as a function of the size of the covariance (described by the scaling factor k). Figure 5.21 depict the P_c curves resulted from three methods with the scaling factor as a parameter for the conjunction cases.

Figure 5.21 shows that in the case of eccentric orbit, the P_c resulted from the conjunction geometries-based explicit expression (denoted by P_{c1} , line with marker ‘ Δ ’) matches well with the result from the general method (denoted by P_{c0} , line with marker ‘ \circ ’), these lines maybe illegible because they are very close. While the NTW components-based explicit expression (denoted by P_{c2} , line with marker ‘ \square ’) is separated from the result from the general slightly.

Table 5.8 and Fig. 5.21 reveal that the error of the conjunction geometries-based explicit expression is larger than the NTW components-based explicit expression.

Table 5.8 Collision probabilities from Chan’s method and explicit expressions (ISS event)

General method	Explicit expression (conjunction geometries)	Explicit expression (NTW components)
5.080119×10^{-5}	4.749411×10^{-5}	5.097559×10^{-5}

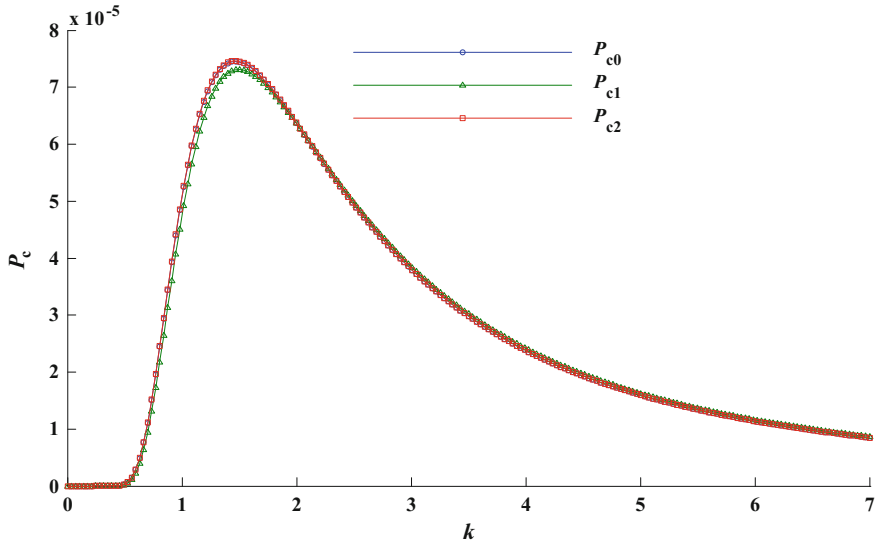


Fig. 5.21 P_c versus error scaling factor using three methods (ISS event)

The reason is that during the derivation of the conjunction geometries-based explicit expression the assumption that the N axis is parallel to vector $\Delta \mathbf{r}_{\min}$ was introduced. The precisions of two explicit expressions are sufficient for the conjunction assessment (the relative errors are less than 7%).

5.4 Explicit Expression’s Eccentricity Bound

The explicit expression of P_c derived from circular orbit assumption is simple and compact. The explicit expression of P_c in terms of relative positions in RSW system for circular orbit is simple and compact, reveals the relationship between P_c and conjunction geometry more directly. But the circular orbit assumption restricts its application. We will determine the bounds of eccentricity for a near-circular orbit to which the explicit expression is applicable.

5.4.1 Flight-Path Angle and Velocity’s Magnitude of Eccentric Orbit

The two major inferences from circular orbit assumption which contribute the derivation of explicit expression are null flight-path angles and equal magnitudes of

velocities. For an eccentric orbit, the expression for the flight-path angle θ and its maximum value θ_{\max} are [51]

$$\begin{cases} \tan \theta = \frac{e \sin f}{1 + e \cos f} \\ \theta_{\max} = \pm \arcsin e, \quad \cos f = -e, \quad f = \pm \left(\frac{\pi}{2} + \sin^{-1} e \right) \end{cases} \quad (5.80)$$

where e is the eccentricity and f is the true anomaly. Equation (5.80) provides the relationship between the flight-path angle and the eccentricity and true anomaly.

The dimensionless parameter η is defined as the ratio of the velocity's magnitude v of the eccentric orbit to the velocity's magnitude v_c of circular orbit with the semimajor equaling to the radius. The ratio η and its maximum/minimum values are

$$\begin{aligned} \eta^2 &= \frac{1 + 2e \cos f + e^2}{1 + e \cos f} \\ \eta_{\max} &= \sqrt{1 + e}, \quad f = 0 \\ \eta_{\min} &= \sqrt{1 - e}, \quad f = \pi \end{aligned} \quad (5.81)$$

Figure 5.22 shows various values for the flight-path angle θ and the ratio η with respect to true anomaly f when $e = 0.7$. An interesting and useful property could be obtained from Fig. 5.22. The flight-path angle θ equals zero when the ratio η is maximum or minimum (perigee or apogee), while the flight-path angle θ achieves its maximum value when the ratio η equals to one. This property indicates that the effects of non-zero flight-path angle and unequal velocity's magnitude on the

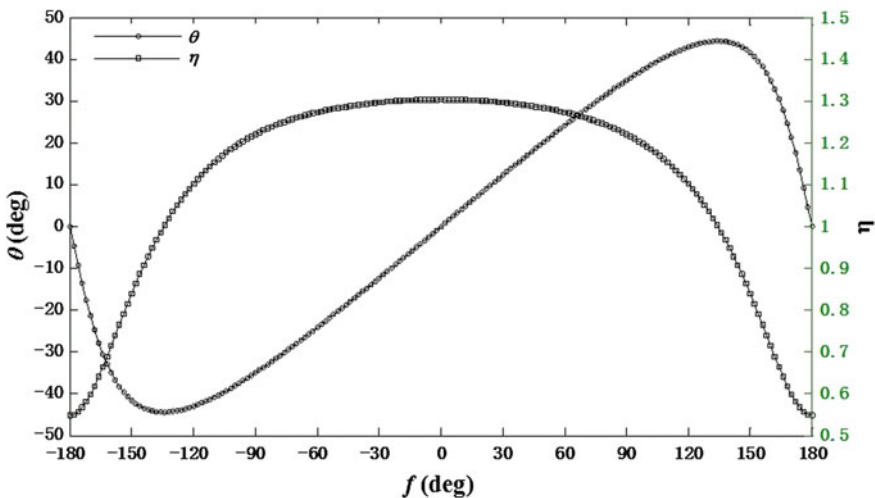


Fig. 5.22 Values for the flight-path angle θ and the ratio of velocity's magnitude η with respect to true anomaly when $e = 0.7$

explicit expression of P_c are not synchronous. One effect tends to maximum while another tends to zero. This property is convenient for the analysis.

The violations of the null flight-path angle and equal velocity’s magnitude are the explicit expression’s major error source, when used to the eccentric orbit. Considering the property of flight-path angle and velocity’s magnitude in eccentric orbit, we can gain an insight into the bounds of eccentricity regarding the explicit expression.

In this section the error of the explicit expression introduced by the eccentric orbit will be demonstrated. The position coordinates of objects in conjunction are described in Table 5.1, the positional error standard deviations in their NTW systems are given in Table 5.3. We consider them as in the RSW system when use the explicit expression Eq. (5.26) although the errors are given in the NTW system. We suppose that the position coordinates of objects in conjunction are the values listed in Table 5.2. However, the flight-path angles and magnitudes of velocities of secondary objects are variable.

Based on the property that the effects of non-zero flight-path angle and unequal velocity’s magnitude are not synchronous, we will first analyze them individually using the U.S. and Russian satellites’ collision event, and then combine them to get the general effect using both cases.

5.4.2 Effects of Non-zero Flight-Path and Unequal Velocity

The flight-path angle is zero at perigee and apogee of the eccentric orbit, the effect of unequal velocities’ magnitudes achieves its maximum. The ratios η at perigee and apogee are $\eta_{\text{perigee}} = \sqrt{1+e}$ and $\eta_{\text{apogee}} = \sqrt{1-e}$, respectively. We suppose that the primary orbit is circular and the secondary orbit is eccentric, the eccentricity of secondary object varies from zero to 0.7. Figure 5.23 shows the P_c calculated by explicit expression (line with marker “o”) and modified expression (line with marker “Δ”) if the conjunction occurs at the perigee or apogee of secondary orbit.

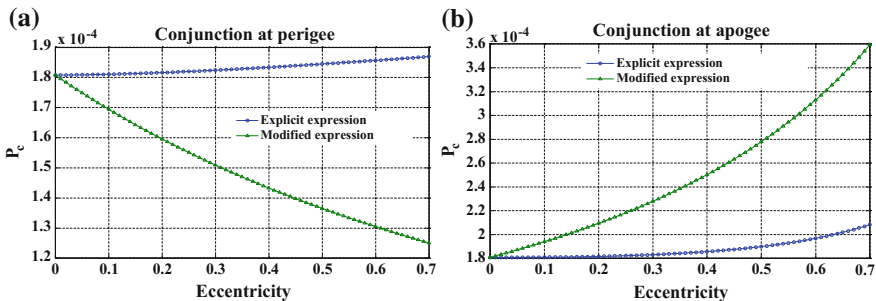


Fig. 5.23 The P_c calculated by explicit expression and modified expression if the conjunction occurs at the perigee (a) and apogee (b) of secondary orbit

In eccentric orbit the ratio $\eta = 1$ when the flight-path angle achieves its maximum value $\theta_{\max} = \pm \sin^{-1} e$. We suppose that the primary orbit is circular and the secondary orbit is eccentric, the eccentricity of secondary object varies from zero to 0.7. Figure 5.24 shows the P_c calculated by explicit expression (line with marker “o”) and modified expression (line with marker “Δ”) if the conjunction occurs at the position on secondary orbit where $\theta = \arcsin e$ or $\theta = -\arcsin e$.

Figures 5.23 and 5.24 indicate that the error introduced by non-zero flight-path angle is much larger than the error introduced by unequal velocity’s magnitude. The non-zero flight-path angle is the dominant effect derived from eccentric orbit while the unequal velocity’s magnitude is the minor one.

5.4.3 General Effects of Eccentric Orbit

We also suppose that the primary orbit is circular and the secondary orbit is eccentric, the eccentricity of secondary orbits cover a series of typical values (0.001, 0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, and 0.7). The flight-path angle and ratio of velocity’s magnitude according to various true anomalies are calculated by Eqs. (5.80) and (5.81). Figure 5.25 shows the differences between P_c calculated by explicit expression and modified expression with various true anomalies. Figure 5.25 indicates that the absolute difference ΔP_c in U.S. and Russian satellite collision event is about two orders of magnitude larger than in ISS event. Considering that the collision risk associated with U.S. and Russian satellite collision event is about two orders of magnitude larger than the one associated with ISS event (the P_c are 1.8×10^{-4} and 8.3×10^{-7}), the relative differences of P_c in two cases are of the same order of magnitude.

As was mentioned previously, the flight-path angle is zero and the ratio of velocity’s magnitude is maximum or minimum at perigee or apogee. Because the non-zero flight-path angle is the dominant effect of error, the error is relative small near the perigee and apogee, as shown in Fig. 5.25. However, it should be noted that the position where the error is minimum is not exactly at perigee or apogee.

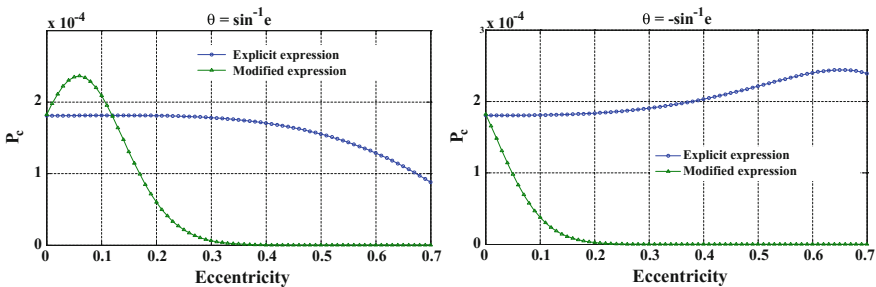


Fig. 5.24 The P_c calculated by explicit expression and modified expression if the conjunction occurs at the position on secondary orbit where $\theta = \pm \sin^{-1} e$

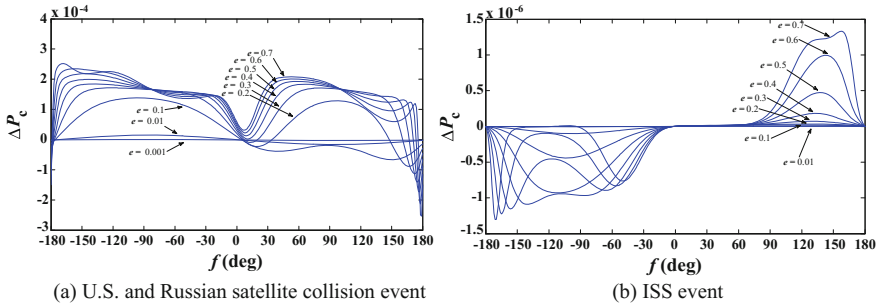


Fig. 5.25 The differences between P_c calculated by explicit expression and modified expression with various true anomalies

Moreover, for most part of the orbit the smaller the eccentricity, the smaller the error is.

LEO is the most crowded region in space and also the key region the conjunction assessment handling with. The explicit expression of P_c is mainly used to process the conjunction events in this region. According to the space surveillance data provided by the website www.Space-Track.org which is maintained by U.S. Strategic Command (USSTRATCOM), the total number of man-made objects in orbit is 16,686 as of January 2013, of which 12,043 objects are in LEO (altitude of apogee less than 5000 km). Figure 5.26 plots the cumulative distribution function (CDF) of the eccentricity of these 12,043 LEO objects. Figure 5.26 indicates that the objects with eccentricity less than 0.1 account for 98.8% of all LEO objects, and the objects with eccentricity less than 0.03, 0.02, and 0.01 account for 90, 82, and 62% of all LEO objects, respectively.

Figure 5.27 shows the relative error between P_c calculated by explicit expression and modified expression with various true anomalies when the eccentricity of secondary object is 0.03, 0.02, 0.01, 0.001.

Figure 5.27 indicates that for objects with eccentricity less than 0.01 (which account for 62% of all LEO objects), the relative error of the explicit expression is less than 0.107 for U.S. and Russian satellite collision event and less than 0.252 for ISS event. The more detailed results of relative error for two conjunction cases in various eccentric orbits are summarized in Table 5.9. Figure 5.27 reveal that for high risk conjunction (U.S. and Russian satellite collision event with $P_c = 1.8 \times 10^{-4}$) the relative error is smaller while for low risk conjunction (ISS event with $P_c = 8.3 \times 10^{-7}$) the relative error is larger. This property validates the application of explicit expression in practical conjunction assessment.

What really matters in conjunction risk assessment is the order of magnitude rather than the specific value of P_c . For example, the relative error of a P_c valued 2×10^{-6} to P_c valued 1×10^{-6} is 100%, but the collision risk level associated with these two values are pretty much the same. Therefore the precision of explicit expression is sufficient for conjunction risk assessment and decision-making for most LEO objects.

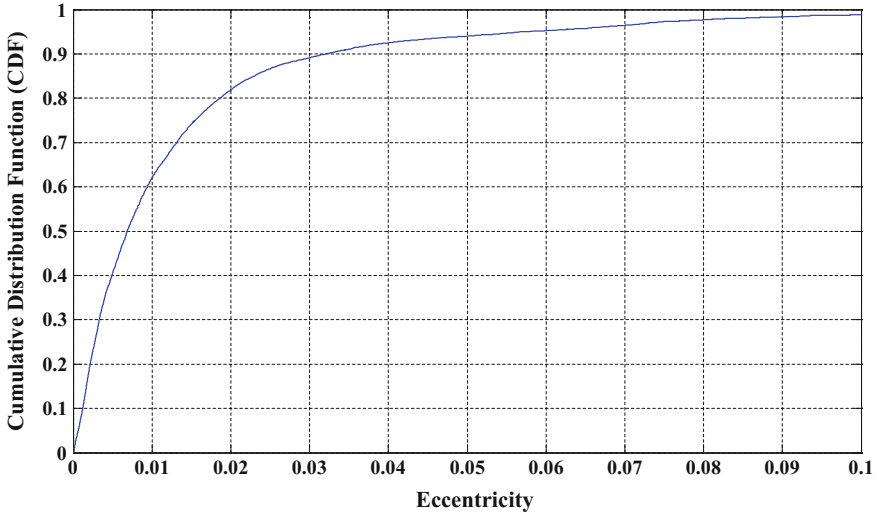


Fig. 5.26 Cumulative distribution function of eccentricity of 12,043 objects in LEO ($h_a < 5000$ km)

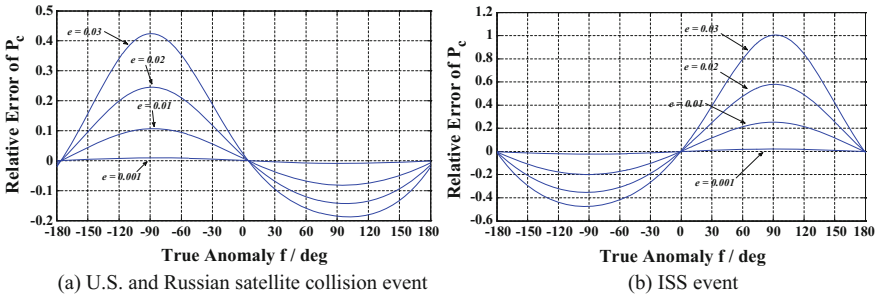


Fig. 5.27 The relative error between P_c calculated by explicit expression and modified expression with various true anomalies

Table 5.9 Relative error for two cases in various eccentric orbits

e	Percent of LEO objects (%)	Relative error (U.S. and Russian satellite event)	Relative error (ISS event)
<0.03	90	0.425	1.006
<0.02	82	0.246	0.579
<0.01	62	0.107	0.252

The explicit and modified expressions of P_c have improved the collision probability analysis in conjunction risk assessment, the benefits is that the explicit expression of P_c relates the collision probability and components of relative

position which are two kinds of important conjunction risk criteria. It is convenient to investigate the effects of influencing factors such as conjunction geometry and position error covariance on P_c and the sensitivity of P_c to the conjunction geometry and covariance.

Chan's analytical formulae of collision probability is the foundation of this work, so the restriction that the collision cross-section dimensions should be small compared to the standard deviations or the miss distance is also applicable to the explicit expression. Moreover, influencing factors and sensitivity analysis of collision probability based on explicit expression need to be done. This study is under progress now.

References

1. Chan FK (2004) Short-term vs. long-term spacecraft encounters. AIAA 2004-5460. AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Providence, Rhode Island
2. Foster JL, Estes HS (1992) A parametric analysis of orbital debris collision probability and maneuver rate for space vehicles. NASA/JSC-25898. NASA Johnson Space Flight Center, Houston
3. Chan FK (2008) Spacecraft collision probability. The Aerospace Press, El Segundo
4. Chan FK (1997) Collision probability analyses for Earth-orbiting satellites. *Adv Astronaut Sci* 96:1033–1048
5. Chan FK (2003) Improved analytical expressions for computing spacecraft collision probabilities. AAS 03-184. AAS/AIAA Space Flight Mechanics Meeting, Ponce, Puerto Rico
6. Alfano S (2004) Accommodating rectangular objects in probability calculations. AIAA 2004-5217. AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Providence, Rhode Island
7. Alfano S (2005) A numerical implementation of spherical object collision probability. *J Astronaut Sci* 53(1):103–109
8. Alfano S (2006) Satellite collision probability enhancements. *J Guidance Control Dyn* 29(3):588–592
9. Patera RP (2001) General method for calculating satellite collision probability. *J Guidance Control Dyn* 24(4):716–722
10. Patera RP (2004) Conventional form of the collision probability integral for arbitrary space vehicle shape. AIAA 2004-5218. AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Providence, Rhode Island
11. Patera RP (2005) Calculating collision probability for arbitrary space-vehicle shapes via numerical quadrature. *J Guidance Control Dyn* 28(6):1326–1328
12. Patera RP (2002) Quick method to determine long-term orbital collision risk. AIAA 2002-1809. Sat-Max2002—Satellite Performance Workshop, Arlington, VA
13. Alfano S (2007) Beta conjunction analysis tool. AAS 07 393. AAS/AIAA Space Flight Mechanics Meeting, Sedona, Arizona, 28 Jan–01 Feb 2007
14. Chan FK (2002) Determination of minimum spacecraft separation at conjunction. AIAA/AAS Astrodynamics Specialist Conference, Monterey, CA
15. Alfano S (2007) Review of conjunction probability methods for short-term encounters. AAS 07-148. AAS/AIAA Space Flight Mechanics Meeting, Sedona, Arizona, 28 Jan–01 Feb 2007
16. CCSDS Secretariat (2012) Draft recommendation for space data system standards: conjunction data message. Draft recommended standard CCSDS 508.0-R-1
17. Wang H (2007) Control and trajectory safety of rendezvous and docking. National University of Defense Technology, Changsha

18. Wang H, Li HY, Tang GJ (2006) General method for calculating spacecraft collision probability. *J Natl Univ Defense Technol* 28(4):27–31
19. Wang H, Li HY, Tang GJ (2007) The V-bar departure control in rendezvous and docking under field of view constraint. *J Astronaut* 28(1):28–32
20. Wang H, Li HY, Tang GJ (2007) Collision probability based trajectory safety in close range guidance phase of rendezvous and docking. *J Astronaut* 28(3):648–652
21. Wang H, Li HY, Tang GJ (2007) Trajectory safety for close range guidance phase of rendezvous and docking. *J Astronaut* 28(6):648–652
22. Wang H, Li HY, Tang GJ (2008) Collision probability based optimal collision avoidance maneuver in rendezvous and docking. *J Astronaut* 29(1):220–223
23. Cheng T, Liu J, Wang RL et al (2006) Collision probability analysis and application of cataloged space debris. *Chin J Space Sci* 26(6):452–458
24. Cheng T (2006) Collision probability analysis and application of cataloged space debris. Graduate School of Chinese Academy of Sciences, Beijing
25. Feng H (2008) Analysis and research for space debris collision probability and its threshold. Graduate School of Chinese Academy of Sciences, Beijing
26. Yang X (2010) Analysis and research for space debris collision probability and its sensitivity. Graduate School of Chinese Academy of Sciences, Beijing
27. Zhang MX (2010) Research on calculation method of spacecraft collision probability. Harbin Institute of Technology, Harbin
28. Wu B (2011) Research on collision probability between space objects during encounter. The PLA Information Engineering University, Zhengzhou
29. Chan FK (2003) Spacecraft collision probability for long-term encounters. AAS 03-549. AAS/AIAA Astrodynamics Specialist Conference, Big Sky, Montana, 3–7 Aug 2003
30. Patera RP (2003) Satellite collision probability for non-linear relative motion. *J Guidance Control Dyn* 26(5):728–733
31. Patera RP (2006) Collision probability for larger bodies having nonlinear relative motion. *J Guidance Control Dyn* 29(6):1468–1471
32. Slater GL, Byram SM, Williams TW (2006) Collision avoidance for satellites in formation flight. *J Guidance Control Dyn* 29(5):1139–1146
33. Alfano S (2006) Addressing nonlinear relative motion for spacecraft collision probability. AIAA 2006-6760. AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Keystone, Colorado
34. McKinley DP (2006) Development of a nonlinear probability of collision tool for the Earth observing system. AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Keystone, Colorado
35. Xu XL, Xiong YQ (2011) A research on collision probability calculation of space debris for nonlinear relative motion. *Acta Astron Sinica* 52(1):73–85
36. Zhang G (2008) Collision prediction and avoidance control for satellites in formation flying. Dalian University of Technology, Dalian, p 6
37. Luo YZ, Liang LB, Wang H, Tang GJ (2011) Quantitative performance for spacecraft rendezvous trajectory safety. *J Guidance Control Dyn* 34(4):1264–1269
38. Liang LB, Luo YZ, Tang GJ, Li HY (2010) Safety-optimal impulsive rendezvous with trajectory uncertainties. In: 61st international astronomical congress, Prague, Czech Republic
39. Liang LB, Luo YZ, Wang H et al (2010) Characteristic analysis of space rendezvous trajectory safety. *J Natl Univ Defense Technol* 32(5):17–22
40. Liang LB (2011) Quantitative performance and design optimization approach of close-range rendezvous trajectory safety. National University of Defense Technology, Changsha, p 4
41. Liang LB, Luo YZ, Wang H et al (2010) Study on the quantitative performance index of space rendezvous trajectory safety. *J Astronaut* 31(10):2239–2245
42. Coppola VT, Woodburn J, Hujsak R (2004) Effects of cross correlated covariance on spacecraft collision probability. AAS 04-181. The AAS/AIAA Space Flight Mechanics Meeting

43. Bird D (2013) Conjunction summary message guide. 64th international astronomical congress, Beijing, China
44. Alfriend KT, Akella MR, Frisbee J et al (1999) Probability of collision error analysis. *Space Debris* 1:21–35
45. Bai XZ, Chen L (2008) Research on calculational method of collision probability between space objects. *J Astronaut* 29(4):1435–1442
46. Bai XZ, Chen L (2009) A rapid algorithm of space debris collision probability based on space compression and infinite series. *Acta Math Applicatae Sinica* 32(2):336–353
47. Zheng QY, Wu LD (2004) A computation method to warn the collision event between space probe and debris. *Acta Astron Sinica* 45(4):422–427
48. Liu J, Wang RL, Zhang HB (2004) Space debris collision prediction research. *Chin J Space Sci* 24(6):462–469
49. Hoots FR, Crawford LL, Roehrich RL (1984) An analytic method to determine future close approaches between satellites. *Celest Mech* 33(2):143–158
50. Liou JC (2009) ISS crew seeks safe haven during debris flyby. *Orbital Debris Q News* 13(4):3
51. Vallado DA (2004) *Fundamentals of astrodynamics and applications*, 2nd edn. Microcosm Press, El Segundo

Chapter 6

Application of Collision Probability

6.1 Sensitivity Analysis of Collision Probability

Collision probability analysis includes collision probability calculation and sensitivity analysis of collision probability [1]. Given that the parameters needed for computing the P_c are uncertain, one should be cautious in making a decision based upon a P_c computed using nominal values. It would be prudent to determine the sensitivity of the computed probability to the expected variations in the parameters describing the computation [2]. The effects of influencing factors (conjunction geometry, position error covariance, object's size) on P_c calculations are investigated in the sensitivity analysis of P_c .

The sensitivity of P_c to conjunction geometry is significant for an optimal strategy for making an avoidance maneuver and analyzing the effect of bias of orbital prediction on the P_c calculation. The P_c calculation is most sensitive to the quality of the covariance. The improper covariance will trigger a false alarm or miss alarm. The sensitivity of P_c to covariance is significant for maximum P_c estimation, false alarm and missing alarm analysis, and probability dilution analysis.

The sensitivity analysis is divided into two basic methods: analytical method and numerical method. The analytical method aims to calculate the partial derivative of independent variables based on the analytical expression of dependent variables so as to obtain the analytical expression of sensitivity. The objective of the numerical method is to alter the independent variable value to make the dependent variable value change and to obtain the sensitivity for some complicated functional relations that are difficult to work out their analytical expressions. Considering the difficulties of getting the analytical expression of sensitivity with the numerical method, the changing rule of sensitivity hence is unobtainable.

An analysis has been done on effects of orbital data quality upon the feasibility of collision risk management by Jenkin [3]. Newman and Duncan integrated collision probability sensitivity analysis into 2D collision probability utility, which consists of the miss distance, various collision probability calculations, and

conjunction geometry information [1]. Frigm et al. plotted the P_c sensitivity surface versus a combined uncertainty scale factor and the hard body radius, which graphically shows the change in P_c due to variations in hard body radius and combined position uncertainty [4]. Alfriend et al. analyzed the sensitivities of collision probability to the covariance size, conjunction distance, orbital plane included angle and other factors [5] by drawing a surface with a collision probability-based one-dimensional integral method. However, the two- or three-dimensional plots do not completely characterize the variation of P_c to variations in each collision parameters. Moreover, the P_c surface can only provide a qualitative insight into the sensitivity. It is inconvenient to obtain the quantitative assessment of sensitivity.

The explicit expression of collision probability bridges the collision probability with conjunction parameters and error parameters. The explicit expression of P_c in terms of collision parameters is a powerful tool for sensitivity analysis. With it, the sensitivities of collision probability to the conjunction geometry, orbital error covariance and object size can be analytically studied. Orbit prediction errors are principal influential factors of collision probability, and further studies can be conducted on how the former affects are like based on sensitivity analysis.

6.1.1 Definition of Sensitivity of Collision Probability

The sensitivity represents the extent of acute changes of dependent variables caused by the changes of independent variables. An n -ary function was set as below:

$$y = f(x_1, x_2, \dots, x_n) \quad (6.1)$$

n refers to the total number of independent variable and y means the function value taken from a certain point within its definitional domain. Δx_i is an indicator of a minor change in the i th independent variable x_i while Δy signifies the changes in corresponding function value.

The sensitivity of the dependent variable y to the independent variable x_i can be defined in two ways:

- (1) Sensitivity S_{1x_i} was defined as the ratio between the change value of dependent variables and independent variables as per their respective absolute variation, viz.,

$$S_{1x_i} = \frac{\Delta y}{\Delta x_i} \quad (6.2)$$

A partial derivative was substituted for the derivative when the change value approximated to 0 to get the following:

$$S_{1x_i}(x_1, \dots, x_n) = \frac{\partial y}{\partial x_i} \quad (6.3)$$

- (2) Sensitivity S_{2x_i} was defined as the ratio between the relative change rates of dependent variables and independent variables based on their respective relative change variation, viz.,

$$S_{2x_i} = \frac{\Delta y/y}{\Delta x_i/x_i} \quad (6.4)$$

A partial derivative was substituted for the derivative when the change value was close to 0 to get the following:

$$S_{2x_i}(x_1, \dots, x_n) = \frac{\partial y/y}{\partial x_i/x_i} = \frac{x_i}{y} \cdot \frac{\partial y}{\partial x_i} \quad (6.5)$$

Sensitivity S_{2x_i} is a dimensionless quantity. Like $f(\cdot)$, the two sensitivities S_{1x_i} and S_{2x_i} defined above are both functions of independent variables (x_1, x_2, \dots, x_n) , indicating the extent of acute changes caused to dependent variables by the changes in a certain independent variable on this point within the domain of definition.

6.1.2 Sensitivity Analysis Based on Explicit Expression

Under the conditions of near-circular orbits and general orbits, the explicit expression of collision probability in terms of conjunction geometry and conjunction distance component is obtained and is explained in Chap. 5. Under the condition of circular orbits, the explicit expressions represented by the components R , S and W of conjunction distance are simple and univocal and easy to be analyzed. In this section, they are used as examples to analyze the sensitivity. Other forms of explicit expressions are not to be explained in details again.

The explicit expression of collision probability in terms of the R , S and W components of miss distance is

$$\begin{aligned} P_c &= P_c(R, S, W, \varphi, \sigma_R, \sigma_S, \sigma_W, r_A) \\ &= \exp \left[-\frac{1}{2} \left(\frac{R^2}{\sigma_R^2} + \frac{S^2 + W^2}{\sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2}} \right) \right] \cdot \left[1 - \exp \left(-\frac{r_A^2}{2\sigma_R \sqrt{\sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2}}} \right) \right] \end{aligned} \quad (6.6)$$

The collision probability was written in the simplified form below to ease the derivation:

$$P_c = \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right] \cdot \left[1 - \exp \left(-\frac{r_A^2}{2\sigma_x\sigma_y} \right) \right] \tag{6.7}$$

where

$$\begin{cases} x = R, & y = \sqrt{S^2 + W^2} \\ \sigma_x^2 = \sigma_R^2, & \sigma_y^2 = \sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2} \end{cases}$$

Below are the sensitivities of collision probability to the components R, S and W of conjunction distance, conjunction angle φ , position error standard deviation $\sigma_{R,S,W}$, and equivalent radius r_A deduced from Eq. (6.7).

1. Sensitivities to R, S and W components

In the Eq. (6.7), the partial derivatives of P_c to R, S and W are taken to obtain the following equations:

$$\begin{cases} \frac{\partial P_c}{\partial R} = \frac{\partial P_c}{\partial x} \frac{\partial x}{\partial R} = -P_c \frac{R}{\sigma_x^2} \\ \frac{\partial P_c}{\partial S} = \frac{\partial P_c}{\partial y} \frac{\partial y}{\partial S} = -P_c \frac{S}{\sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2}} \\ \frac{\partial P_c}{\partial W} = \frac{\partial P_c}{\partial y} \frac{\partial y}{\partial W} = -P_c \frac{W}{\sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2}} \end{cases} \tag{6.8}$$

The two sensitivities of collision probability P_c to R, S and W are listed in Table 6.1.

The minus “-” in Table 6.1 implies that the collision probability P_c would decrease monotonically with the increase of R, S and W.

2. Sensitivities to orbital error standard deviations

The partial derivatives of functions $P_c = P_c(\sigma_x, \sigma_y)$ with respect to σ_x and σ_y are respectively expressed as below:

$$\begin{cases} \frac{\partial P_c}{\partial \sigma_x} = e^{-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right)} \left(1 - e^{-\frac{r_A^2}{2\sigma_x\sigma_y}} \right) \frac{x^2}{\sigma_x^3} - e^{-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right)} e^{-\frac{r_A^2}{2\sigma_x\sigma_y}} \frac{r_A^2}{2\sigma_x^2\sigma_y} \\ \frac{\partial P_c}{\partial \sigma_y} = e^{-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right)} \left(1 - e^{-\frac{r_A^2}{2\sigma_x\sigma_y}} \right) \frac{y^2}{\sigma_y^3} - e^{-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right)} e^{-\frac{r_A^2}{2\sigma_x\sigma_y}} \frac{r_A^2}{2\sigma_x\sigma_y^2} \end{cases} \tag{6.9}$$

Table 6.1 Two types of sensitivities to R, S and W

Sensitivity	R	S	W
S ₁	$-P_c \frac{R}{\sigma_R^2}$	$-P_c \frac{S}{\sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2}}$	$-P_c \frac{W}{\sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2}}$
S ₂	$-\frac{R^2}{\sigma_R^2}$	$-\frac{S^2}{\sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2}}$	$-\frac{W^2}{\sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2}}$

The expression of collision probability is substituted into the Eq. (6.9) to get the followings:

$$\begin{cases} \frac{\partial P_c}{\partial \sigma_x} = \frac{P_c}{\sigma_x} \left(\frac{x^2}{\sigma_x^2} + \frac{r_A^2}{2\sigma_x\sigma_y} \left(1 - \frac{1}{1 - e^{-\frac{r_A^2}{2\sigma_x\sigma_y}}} \right) \right) \\ \frac{\partial P_c}{\partial \sigma_y} = \frac{P_c}{\sigma_y} \left(\frac{y^2}{\sigma_y^2} + \frac{r_A^2}{2\sigma_x\sigma_y} \left(1 - \frac{1}{1 - e^{-\frac{r_A^2}{2\sigma_x\sigma_y}}} \right) \right) \end{cases} \quad (6.10)$$

If the dimensionless variables A_R and A_{SW} are described in the equation below:

$$\begin{cases} A_R = \frac{x^2}{\sigma_x^2} + \frac{r_A^2}{2\sigma_x\sigma_y} \left(1 - \frac{1}{1 - e^{-\frac{r_A^2}{2\sigma_x\sigma_y}}} \right) \\ A_{SW} = \frac{y^2}{\sigma_y^2} + \frac{r_A^2}{2\sigma_x\sigma_y} \left(1 - \frac{1}{1 - e^{-\frac{r_A^2}{2\sigma_x\sigma_y}}} \right) \end{cases} \quad (6.11)$$

The sensitivity to the combined error standard deviation σ_R in the R direction can be expressed as follows:

$$S_{1\sigma_R} = \frac{P_c}{\sigma_R} A_R, \quad S_{2\sigma_R} = A_R \quad (6.12)$$

Below the sensitivity of P_c is derived to the standard deviations σ_S and σ_W in the directions of S and W. From Eq. (6.7), we can get the followings:

$$\begin{cases} \frac{\partial \sigma_y}{\partial \sigma_S} = \frac{\sigma_S}{\sigma_y} \cos^2 \frac{\varphi}{2} \\ \frac{\partial \sigma_y}{\partial \sigma_W} = \frac{\sigma_W}{\sigma_y} \sin^2 \frac{\varphi}{2} \end{cases} \quad (6.13)$$

And then the following equation is obtained:

$$\begin{cases} \frac{\partial P_c}{\partial \sigma_S} = \frac{\partial P_c}{\partial \sigma_y} \cdot \frac{\partial \sigma_y}{\partial \sigma_S} = A_{SW} P_c \frac{\sigma_S}{\sigma_y^2} \cos^2 \frac{\varphi}{2} \\ \frac{\partial P_c}{\partial \sigma_W} = \frac{\partial P_c}{\partial \sigma_y} \cdot \frac{\partial \sigma_y}{\partial \sigma_W} = A_{SW} P_c \frac{\sigma_W}{\sigma_y^2} \sin^2 \frac{\varphi}{2} \end{cases} \quad (6.14)$$

Therefore, the sensitivity of P_c can be calculated to the standard deviations of errors in the directions of R, S and W as listed in Table 6.2.

The plus and minus signs for sensitivity to error standard deviation are determined by the plus and minus signs of A_R and A_{SW} . For the conjunction events of

Table 6.2 Two types of sensitivities to standard deviations of errors in directions R, S and W

Sensitivity	σ_R	σ_S	σ_W
S_1	$A_R \frac{P_c}{\sigma_R}$	$A_{SW} P_c \frac{\sigma_S}{\sigma_y^2} \cos^2 \frac{\varphi}{2}$	$A_{SW} P_c \frac{\sigma_W}{\sigma_y^2} \sin^2 \frac{\varphi}{2}$
S_2	A_R	$A_{SW} \frac{\sigma_S^2}{\sigma_y^2} \cos^2 \frac{\varphi}{2}$	$A_{SW} \frac{\sigma_W^2}{\sigma_y^2} \sin^2 \frac{\varphi}{2}$

smaller miss distances, the collision probability will decrease as the covariance increases; for the conjunction events of larger miss distances, the collision probability will increase as the covariance decreases. It will not be probability dilution if the sensitivity is positive; otherwise, it will be probability dilution. If the sensitivity is 0, then it will be maximum collision probability.

3. Sensitivities to conjunction angle

The conjunction angle influences the collision probability by the combined error variance in a horizontal plane. From the Eq. (6.7) above, the partial derivative of σ_y with respect to conjunction angle φ can be calculated:

$$\frac{\partial \sigma_y}{\partial \varphi} = \frac{\sigma_W^2 - \sigma_S^2}{4\sigma_y} \sin \varphi \tag{6.15}$$

Hence, the partial derivative of P_c with respect to conjunction angle φ is calculated below:

$$\frac{\partial P_c}{\partial \varphi} = \frac{\partial P_c}{\partial \sigma_y} \cdot \frac{\partial \sigma_y}{\partial \varphi} = A_{SW} P_c \frac{\sigma_W^2 - \sigma_S^2}{4\sigma_y^2} \sin \varphi \tag{6.16}$$

So, the sensitivity of P_c to conjunction angle φ is obtained as listed in Table 6.3.

The authors have analyzed multiple influencing factors of collision probability based on its explicit expression in Refs. [6, 7]. The conclusion made for the impact of the orbital plane included angle upon collision probability is such that the impact of orbital plane included angle φ upon collision probability varies with different shapes of combined error ellipsoids. The S-direction prediction error of space objects is generally greater than the W-direction prediction error. The collision probability increases progressively as the angle φ widens from 0 to π . If the S-direction prediction error is smaller than the W-direction prediction error, the collision probability decreases progressively as the angle φ widens from 0 to π . The conclusion is imperfect. Below are more comprehensive conclusions that have been drawn in line with the sensitivity expression.

The sensitivity of P_c to conjunction angle φ comprises two terms: The first is dimensionless variable A_{SW} , namely the sensitivity of collision probability to the combined error in the horizontal direction. This parameter, determined by the relative size between the error and the close approach distance, represents whether it is in the state of probability dilution; the second is $(\sigma_W^2 - \sigma_S^2)$, which indicates the relative size between the S-direction and W-direction errors.

Table 6.3 Two types of sensitivities to conjunction angle

Sensitivity	φ
S_1	$A_{SW} P_c \frac{\sigma_W^2 - \sigma_S^2}{4\sigma_y^2} \sin \varphi$
S_2	$A_{SW} \varphi \frac{\sigma_W^2 - \sigma_S^2}{4\sigma_y^2} \sin \varphi$

First of all, the characteristics of joint variances in horizontal plane have been discussed. The joint variance can be written as below:

$$\sigma_{SW}^2 = \sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2} = \sigma_S^2 + (\sigma_W^2 - \sigma_S^2) \sin^2 \frac{\varphi}{2} \quad (6.17)$$

As shown in the Eq. (6.17), for space objects, when the S-direction prediction error is larger than the W-direction prediction error ($\sigma_S^2 > \sigma_W^2$), viz. $\sigma_W^2 - \sigma_S^2 < 0$, the combined error in the horizontal plane σ_{SW}^2 will decrease as the conjunction angle φ enlarges from 0 to π . When the S-direction prediction error is smaller than the W-direction prediction error ($\sigma_S^2 < \sigma_W^2$), viz. $\sigma_W^2 - \sigma_S^2 > 0$, the combined error in the horizontal plane σ_{SW}^2 will increase as the conjunction angle φ enlarges from 0 to π . For a majority of space objects, $\sigma_S^2 > \sigma_W^2$ always happens and thus the combined error in the horizontal plane diminishes as the conjunction angle widens.

Two conditions have been given in Fig. 6.1 to show how the combined error standard deviation in a horizontal plane σ_{SW} varies with the orbital plane included angle φ : The first is $\sigma_S = 100$ (distance unit), $\sigma_W = 20$; the second is $\sigma_S = 20$, $\sigma_W = 100$. As illustrated in Fig. 6.1, when $\sigma_S > \sigma_W$, σ_{SW} becomes smaller as φ enhances; when $\sigma_S < \sigma_W$ (a rare case), σ_{SW} increases with φ .

There are two conditions:

- (1) In common cases. It happens when the S-direction prediction error is larger than the W-direction prediction error and the combined error in horizontal plane will reduce as the conjunction angle enlarges. Therefore, if the error in the horizontal plane is not in the state of probability dilution (i.e. $A_{SW} > 0$), the combined error variances in the horizontal plane will fall off as the conjunction angle φ goes up, so as the collision probability, and the sensitivity of P_c to φ will become minus. If the error is in the state of probability dilution

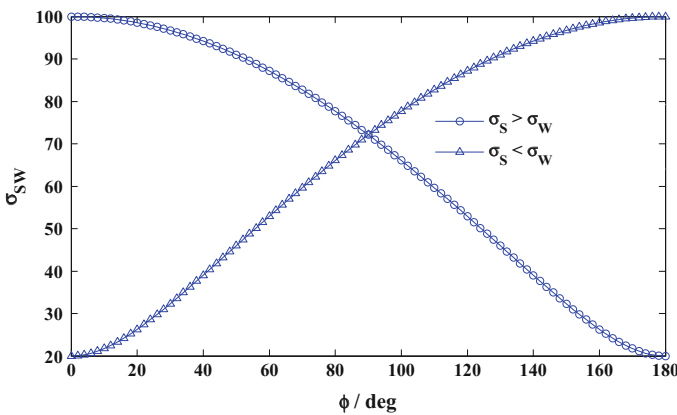


Fig. 6.1 Curve of combined errors in a horizontal plane varying with orbital plane included angles under two conditions

- (i.e. $A_{SW} < 0$), the combined error variance will decrease as φ increases, so as the collision probability, and the sensitivity of which to φ will become plus.
- (2) Under abnormal conditions. It happens when the S-direction prediction error is smaller than the W-direction prediction error and the combined error in a horizontal plane will increase with the conjunction angle. Therefore, if the error in the horizontal plane is not in the state of probability dilution (i.e. $A_{SW} > 0$), the combined error variance in the horizontal plane will enlarge with the conjunction angle φ , so as the collision probability, and the sensitivity of which to φ will become positive. If the error is in the state of probability dilution (i.e. $A_{SW} < 0$), the combined error variance will increase with φ , but the collision probability will decrease and the sensitivity of which to φ will become negative.

The way to determine the sign of sensitivity of collision probability to conjunction angles have been given in Table 6.4. From the table we can see that the conclusions in Refs. [6, 8] are the special case occurred in the condition of probability dilution.

4. Sensitivities to object sizes

The partial dispersion of P_c with respect to the combined radius r_A of the objects is calculated to get results below:

$$\frac{\partial P_c}{\partial r_A} = e^{-\frac{1}{2}\left(\frac{r_A^2}{\sigma_x^2} + \frac{r_A^2}{\sigma_y^2}\right)} \cdot e^{-\frac{r_A^2}{2\sigma_x\sigma_y}} \cdot \frac{r_A}{\sigma_x\sigma_y} = \frac{P_c r_A}{\sigma_x\sigma_y} \left(\frac{1}{1 - e^{-\frac{r_A^2}{2\sigma_x\sigma_y}}} - 1 \right) \tag{6.18}$$

As $r_A \ll \sigma_x, \sigma_y$ is for common conjunction events, so $r_A^2/(\sigma_x\sigma_y) \rightarrow 0$ is a small amount and the approximate value of Eq. (6.18) is

$$\frac{\partial P_c}{\partial r_A} \approx \frac{P_c}{r_A} \left(2 - \frac{r_A^2}{\sigma_x\sigma_y} \right) \tag{6.19}$$

Therefore, the sensitivities of P_c with respect to the combined radius r_A of the objects are

$$\begin{cases} S_{1r_A} = \frac{P_c}{r_A} \left(2 - \frac{r_A^2}{\sigma_x\sigma_y} \right) \\ S_{2r_A} = 2 - \frac{r_A^2}{\sigma_x\sigma_y} \end{cases} \tag{6.20}$$

Table 6.4 Determination of the sign of sensitivity to conjunction angles

	$\sigma_S^2 > \sigma_W^2$ (common)	$\sigma_S^2 < \sigma_W^2$ (unusual)
$A_{SW} > 0$ (no dilution)	Minus (-)	Plus (+)
$A_{SW} < 0$ (dilution)	Plus (+)	Minus (-)

Table 6.5 Expressions of sensitivities of P_c to various influencing factors

Influencing factor	S_1	S_2
R	$-P_c \frac{R}{\sigma_R^2}$	$-\frac{R^2}{\sigma_R^2}$
S	$-P_c \frac{S}{\sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2}}$	$-\frac{S^2}{\sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2}}$
W	$-P_c \frac{W}{\sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2}}$	$-\frac{W^2}{\sigma_S^2 \cos^2 \frac{\varphi}{2} + \sigma_W^2 \sin^2 \frac{\varphi}{2}}$
σ_R	$A_R \frac{P_c}{\sigma_R}$	A_R
σ_S	$A_{SW} P_c \frac{\sigma_S}{\sigma_y} \cos^2 \frac{\varphi}{2}$	$A_{SW} \frac{\sigma_S^2}{\sigma_y^2} \cos^2 \frac{\varphi}{2}$
σ_W	$A_{SW} P_c \frac{\sigma_W}{\sigma_y} \sin^2 \frac{\varphi}{2}$	$A_{SW} \frac{\sigma_W^2}{\sigma_y^2} \sin^2 \frac{\varphi}{2}$
φ	$A_{SW} P_c \frac{\sigma_W^2 - \sigma_S^2}{4\sigma_y^2} \sin \varphi$	$A_{SW} \varphi \frac{\sigma_W^2 - \sigma_S^2}{4\sigma_y^2} \sin \varphi$
r_A	$\frac{P_c}{r_A} \left(2 - \frac{r_A^2}{\sigma_x \sigma_y} \right)$	$2 - \frac{r_A^2}{\sigma_x \sigma_y}$

As shown in the Eq. (6.20) above, the second type of sensitivity (S_2) is approximate to a constant value 2.

With this, the expressions of sensitivities of P_c to the three components R , S , W of the close approach distance, conjunction angle φ , combined error standard deviation $\sigma_{R,S,W}$ of two objects in the RSW coordinate system, and the combined radius r_A of two objects were obtained according to the explicit expression of P_c represented by R , S and W under the condition of circular orbits. In Table 6.5, we have summarized the expressions of these sensitivities.

6.1.3 Examples

The U.S. and Russian satellite collision was taken as an example to analyze the sensitivity of collision probability. The conjunction geometry (close approach distance, component of relative position, velocity included angle, et al.) and position error standard deviation of two satellites at TCA have been given in Sect. 5.2.4. According to the conjunction geometrical parameters and position error standard deviation, the sensitivities of collision probability to various influencing factors can be calculated by the explicit expressions and sensitivity equations mentioned above.

1. Sensitivities to R, S and W components

To study how the sensitivity changes with the close approach distance in the direction R , other conditions were preset to make the close approach distance in the direction R vary from 0 to 0.2 km to calculate P_c and its sensitivities to R S_{1R} and S_{2R} . See the results in Fig. 6.2. As such, the curve of P_c and its sensitivities changing with S and W can be obtained. See Figs. 6.3 and 6.4 for details.

As shown in Figs. 6.2, 6.3 and 6.4, P_c decreases as the close approach distances in the directions R , S and W increases. In fact, the enlargement of relative distances

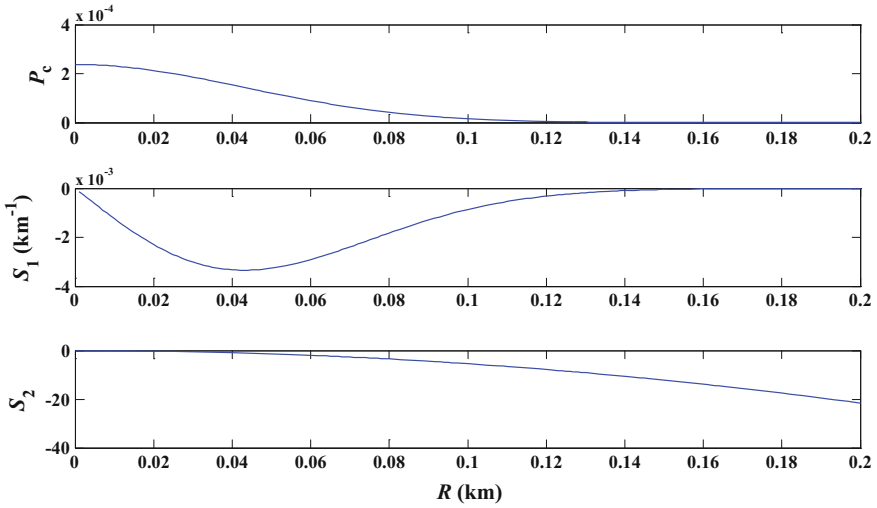


Fig. 6.2 Curve of P_c and its sensitivities changing with R-direction close approach distance

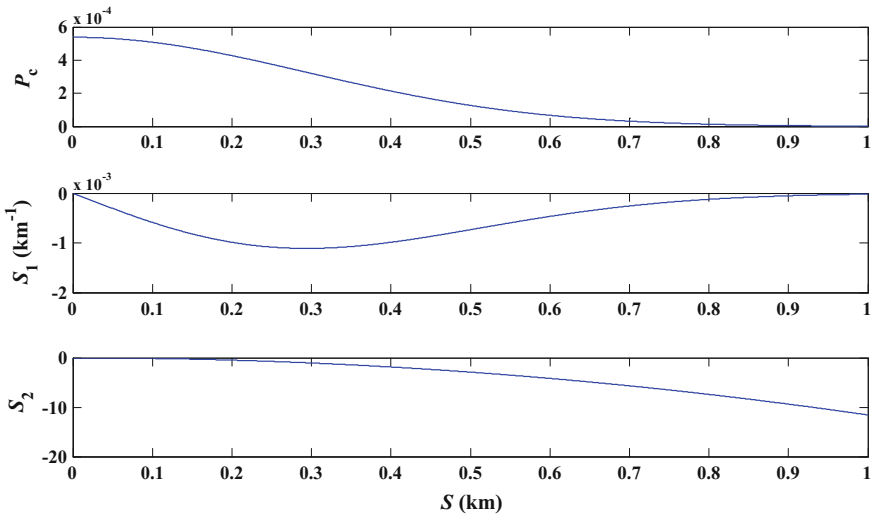


Fig. 6.3 Curve of P_c and its sensitivities changing with S-direction close approach distance

is a major measure taken for spacecraft collision avoidance. The P_c sensitivities to the close approach distances in the directions R, S and W are all constant negatives. As the distance increases from 0, the absolute value of S_1 becomes bigger, then smaller and ultimately close to 0. When the close approach distances in the directions R, S and W are in the bottom of the sensitivity curves in Figs. 6.2, 6.3 and 6.4, the impact of their changes causing P_c to reduce is the greatest.

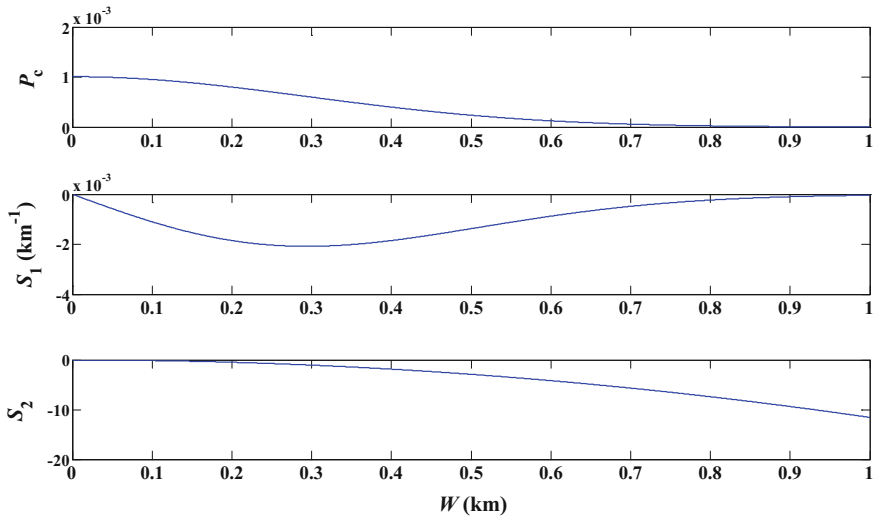


Fig. 6.4 Curve of P_c and its sensitivities changing with W -direction close approach distance

Table 6.6 P_c sensitivities value to components R , S and W

Sensitivity	R	S	W
S_1/km^{-1}	-0.00309429	-0.000908098	-0.00113967
S_2	-0.543051	-2.18255	-3.437644

The P_c sensitivities to the close approach distances in the directions R , S and W under the conjunction parameters and error value conditions described in Sect. 5.2.4 are indicated in Table 6.6.

As shown in Table 6.6, under the conjunction geometry and error conditions of U.S. and Russian satellites, P_c sensitivities to S - and W -direction close approach distances are close to -0.001 in absolute sensitivity S_1 , while that to the R -direction close approach distance is -0.003 . This suggests that it is more reasonable to enlarge the R -direction close approach distance in consideration of reducing the collision probability. But the question of how to do it right in practical operation obviously depends on lots of other factors, such as mission requirement, constraint of impulse applying directions, time to TCA, etc.

2. Sensitivities to orbital error standard deviations

To study the sensitivity changing with the combined error standard deviation, other conditions were preset to make the combined error standard deviation σ_R in the direction R vary from 0 to 0.2 km and the combined error standard deviations σ_S and σ_W in the directions S and W change from 0 to 5 km so as to calculate P_c and its sensitivities to σ_R , σ_S and σ_W . See the result curves in Figs. 6.5, 6.6 and 6.7.

As can be seen in Figs. 6.5, 6.6 and 6.7, P_c increases with the error standard deviations in all the three directions and then decreases when reaching the

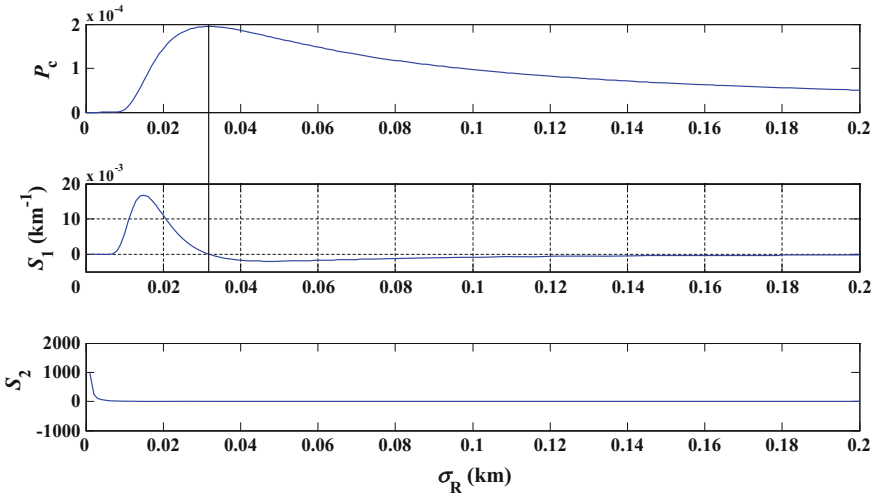


Fig. 6.5 Curve of P_c and its sensitivities changing with the combined error standard deviation in the direction R

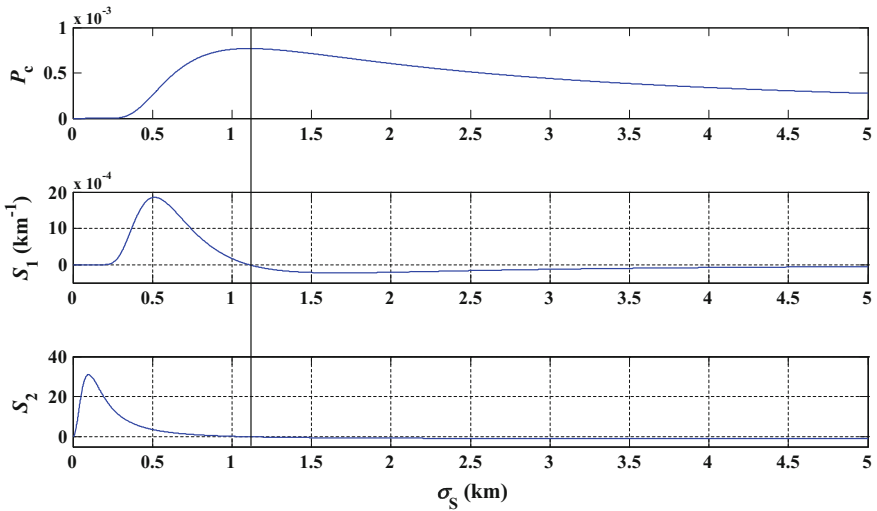


Fig. 6.6 Curve of P_c and its sensitivities changing with the combined error standard deviation in the direction S

maximum value. The same applies to the sensitivity value, i.e., first positive and then negative. After P_c comes to the extreme value, the sensitivity curve passes zero from positive to negative, as marked by a vertical black line in the figure. For the plus part of the sensitivity, it first increases and then decreases to zero. When the value comes to an extreme, the enlargement of the error standard deviation will

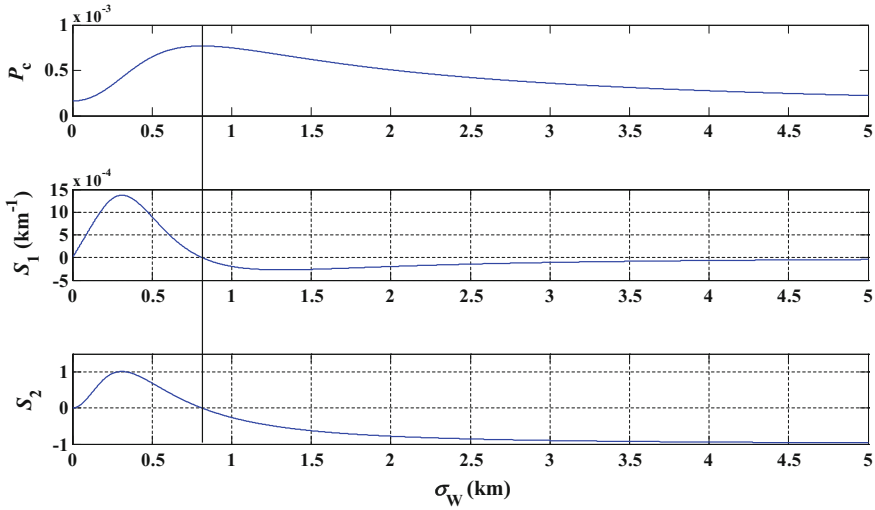


Fig. 6.7 Curve of P_c and its sensitivities changing with the combined error standard deviation in the direction W

Table 6.7 P_c sensitivity value to combined error standard deviations in directions R, S and W

Sensitivity	σ_R	σ_S	σ_W
S_1 (km^{-1})	-0.00191042	0.00173905	0.000467691
S_2	-0.454976	4.416125	0.206047

make P_c enhance at the fastest pace. For the negative part of the sensitivity, the absolute value increases to the extreme and then slowly decreases to zero. The positive part is what is called the area of probability dilution. At this time, the P_c is decreasing as the error increases. The sensitivity can be used to describe the effect size of probability dilution. The curve of P_c changing with σ_S and σ_W is similar to the sensitivity curve in respect to their shapes, which is determined by the coupling properties of error covariance in horizontal planes.

The P_c sensitivities to the combined error standard deviations in the three directions R, S and W under the conjunction parameters and error value conditions described in Sect. 5.2.4 are indicated in Table 6.7. In the table, the P_c sensitivity to σ_R is negative and those to σ_S and σ_W are positive, which suggests that σ_R has already been in the area of probability dilution. As indicated by the absolute value, the P_c sensitivity to σ_S is larger than to σ_W .

3. Sensitivities to conjunction angles

To study how the sensitivity changes with the conjunction angle φ , other conditions were preset to make φ change from 0° to 180° so as to calculate P_c and its sensitivity and draw up their changing curves as shown in Fig. 6.8.

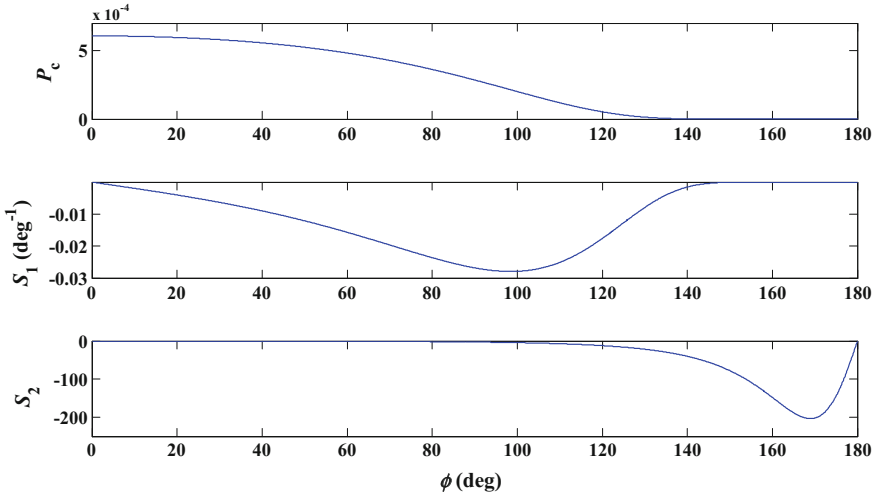


Fig. 6.8 Curve of P_c and its sensitivities changing with the conjunction angles

As is known from the error condition, the error sizes in directions S and W meet the inequation $\sigma_S^2 > \sigma_W^2$. It is common case. The P_c sensitivities to σ_S and σ_W are positive and not in the state of probability dilution. We, therefore, can know from the conclusion in Table 6.4 that the P_c sensitivity is negative, identical to that in Fig. 6.8. Figure 6.8 shows that the current error is small and the combined error in the horizontal plane becomes smaller as the conjunction angle increases and it is always not in the state of probability dilution.

To explain the effects of conjunction angle to P_c when the error becomes bigger, the standard deviations of directional errors have been multiplied by 5 based on the current error. See Fig. 6.9 for the curve of P_c and its sensitivity changing with the conjunction angle. P_c is in the state of probability dilution at first due to large errors. However after that, it increases with the conjunction angle. As the conjunction angle keeps increasing, the combined error in the horizontal plane becomes smaller and smaller to enter the non-probability dilution state. At the same time, P_c reduces as the conjunction angle grows larger. The sensitivity turns from positive to negative.

The standard deviations of errors in different directions have been multiplied by 10 based on current errors. See Fig. 6.10 for the curve of P_c and its sensitivity changing with the conjunction angle. P_c has always been in the state of probability dilution due to large errors. As the conjunction angle increases, the error in the horizontal plane narrows, but P_c increases constantly and its sensitivity to the conjunction angle remains positive.

When the conjunction parameter and error are set as the values in Sect. 5.2.4, the sensitivity of collision probability to conjunction angle is negative, as can be seen in Table 6.8.

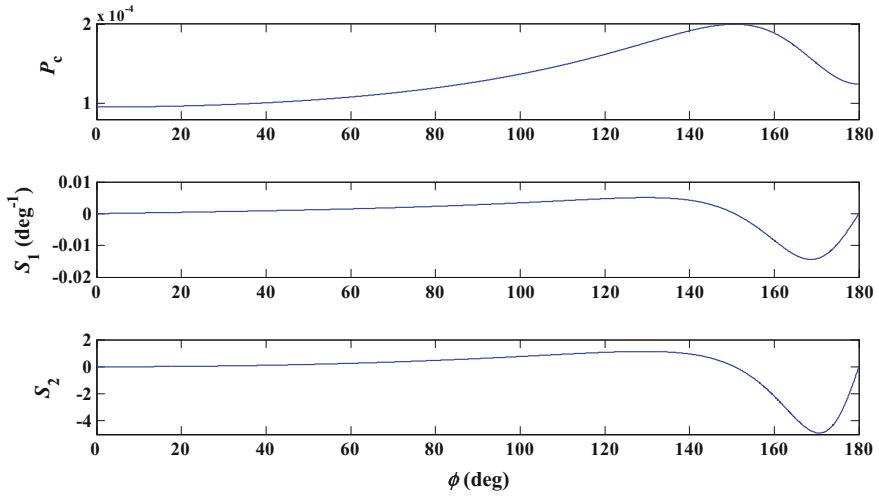


Fig. 6.9 Curve of P_c and its sensitivities changing with the conjunction angles ($\sigma \times 5$)

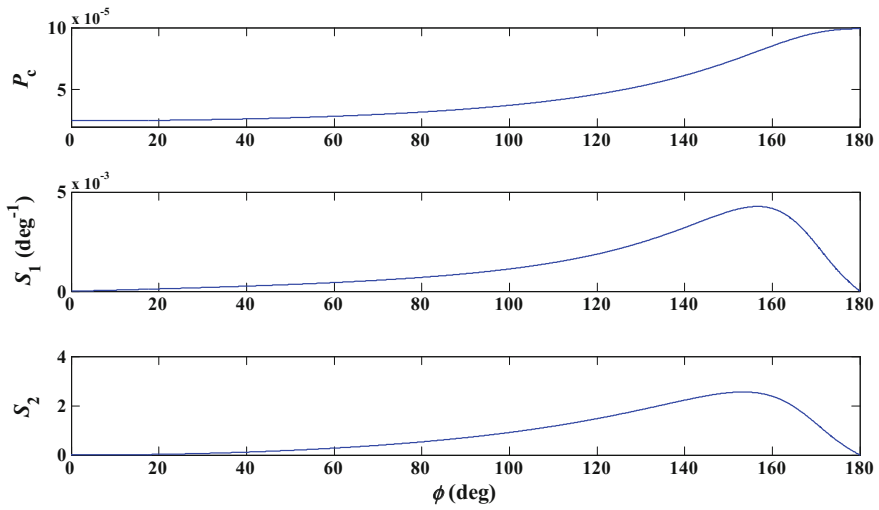


Fig. 6.10 Curve of P_c and its sensitivities changing with the conjunction angles ($\sigma \times 10$)

Table 6.8 P_c sensitivity value to the conjunction angle ϕ

Sensitivity	ϕ
$S_1/(\text{deg}^{-1})^{-1}$	-0.0276207
S_2	-4.768093

4. Sensitivities to object sizes

The applicable conditions for P_c explicit expression are that the object sizes are far smaller than the close approach distances and error standard deviations. Thus, only the smaller objects have been analyzed to get the P_c sensitivity. Other conditions were assumed to be fixed to make the combined radius r_A of two objects change from 0 to 50 m so as to calculate P_c and its sensitivity to object sizes and draw up a changing curve as shown in Fig. 6.11.

P_c increases with the combined radius. Its sensitivity to the object size is always positive and S_2 (relative sensitivity) is around a constant value 2. The figure 2 stands for square relation between combined radius and P_c change. And this is consistent with the results of theoretical analysis stated above.

The P_c sensitivity to the combined radius of two objects under the conjunction parameters and error value conditions described in Sect. 5.2.4 are indicated in Table 6.9.

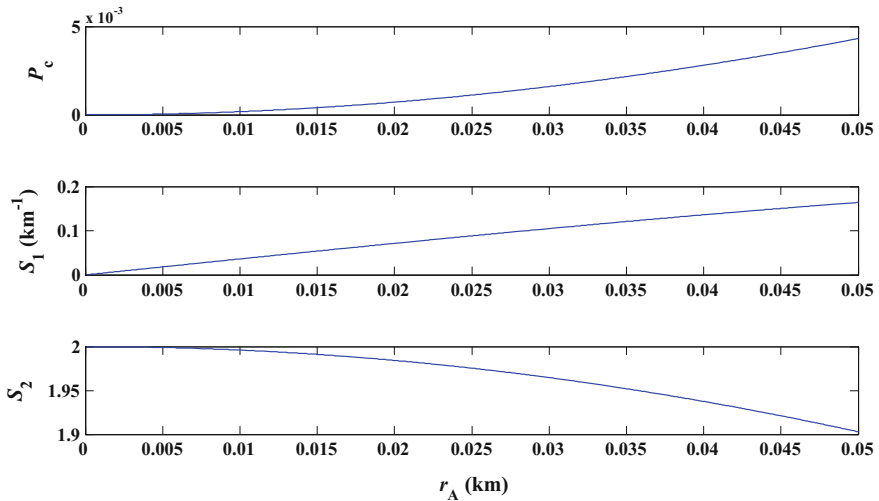


Fig. 6.11 Curve of P_c and its sensitivities changing with combined radius r_A

Table 6.9 P_c sensitivity value to combined radius r_A

Sensitivity	r_A
S_1/km^{-1}	0.0360880
S_2	1.996054

6.2 Maximum Collision Probability

6.2.1 Introduction

For fixed miss distance and space objects' sizes, there is a maximum collision probability (P_{cmax}) that varies with size, shape, and orientation of error covariance ellipsoid. The estimation of P_{cmax} is significant in the conjunction risk assessment. When the covariances of one or both objects are not known or unreliable, P_{cmax} could be useful as an indicator of risk to assess the conjunction in the worst-case.

P_{cmax} could also be used in the pre-filter of dangerous object if the computation time for P_{cmax} is considerably less than the computation of P_c . Actually, as will be seen in this section, P_{cmax} can be computed rapidly by analytical or approximate expressions, while the more time-consuming refined collision probability analysis always includes integral or iteration of series. If P_{cmax} is below the predefined probability threshold, no further refined calculations are needed. As a guideline, the probability threshold should be less than the “yellow” threshold to avoid missing alarm, which is 10^{-5} for the space shuttle. Therefore, this effectively decreases the computational time in using the pre-filter because of the larger number of conjunctions to be considered.

P_{cmax} is normally larger than the real risk. The maximum P_c may not provide a definitive answer to the question of whether or not a risk mitigation maneuver is required. If the calculated maximum P_c is less than the action threshold, no risk mitigation action is necessary, then a useful final result is obtained. However, if the maximum P_c does violate an action threshold, it cannot be concluded that some type of risk mitigation action is absolutely necessary because the actual risk might still be orders of magnitude below that threshold.

P_{cmax} can also determine the distance threshold in the pre-filter of a threatening object. P_{cmax} and corresponding error variance can determine orbital prediction accuracy required for conjunction assessment that will prevent or minimize dilution of the probability calculations [9–11].

Mathematically, an ellipsoid in three-dimension space or an ellipse in two-dimension plane is characterized by its *size*, *shape*, and *orientation*, as illustrated in Fig. 6.12. When the covariance ellipsoids are obtained from independently tracked measurements, the P_c is determined based on the “combined” covariance

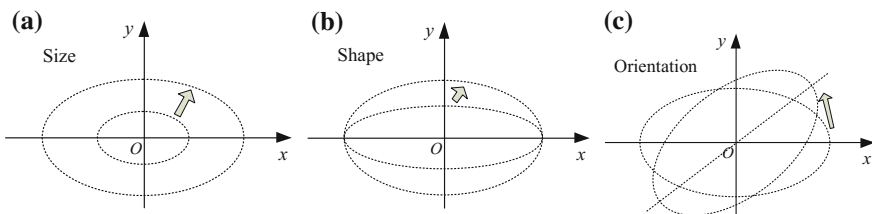


Fig. 6.12 The variation of ellipse's **a** size, **b** shape and **c** orientation in two-dimensions plane

ellipse's size, shape, and orientation associated with each of the objects. The calculated P_c is affected by these three aspects simultaneously. The variation of one or all of them yields a maximum probability. Therefore, when dealing with the maximum collision probability problem it is significant to clarify and understand what assumptions are made to obtain P_{cmax} . Totally different results will be obtained with respect to different definition of P_{cmax} .

In practical conjunction assessment, the predicted orientation of covariance ellipse is generally of high accuracy. That is because the major axis of covariance ellipsoid is usually aligned very closely with the velocity of orbiting object, so the orientation of covariance ellipse in the "encounter" or "conjunction" plane is mainly determined by the conjunction geometry which is often believed to be accurate if the bias of propagation of orbits is ignored. Some previous works treated the orientation as a constant when analyzing the maximum probability [5, 12]. However, the predicted size and shape of covariance ellipse are generally less accurate, especially the former. Almost all works treated the size and shape as variables when handling with the maximum probability.

Alfriend et al. [5] assumed that the probability density function (pdf) is constant over the sphere, and equal to the value at the center of the sphere. Based on this assumption, the value of covariance's scaling factor that maximizes P_c and the corresponding value of P_{cmax} are presented. In this method the size represented by the scaling factor was the only variable, the shape and orientation were fixed.

Alfano [7] developed a method to map regions of maximum probability for various satellite sizes, encounter geometries, and covariance sizes and shapes. Those regions were then examined to assess probability dilution. Charts were created to show the effects of positional uncertainty on probability calculation and assess probability dilution, to determine ephemeris accuracy requirements, and to establish distances for probability-based keep-out zones. In this method the size (1-sigma combined positional deviation), shape (aspect ratio), and orientation were taken into consideration to achieve P_{cmax} . The drawback is that no close-formed solution was given.

Alfano [9] showed how to calculate the upper bounds of probability by determining the "worst" possible covariance parameters and orientation under some extreme assumptions when the aspect ratio of the combined covariance approaches infinity, as well as the major axis of the combined covariance ellipse aligned with the relative position vector. The upper bound of P_c in this extreme situation is often overestimated, which may be too conservative to serve as the discriminator for pre-filtering. As shown in this section, the major axis of the covariance ellipse aligned with the relative position vector is an inevitable result if the variation of orientation is taken into account. This method provides the formulae of P_{cmax} in that extreme situation.

Klinkrad [13] combined two covariances and then scaled the combined covariance to find the maximum pdf for fixed shape and orientation of the error ellipse in the encounter plane. Analogous to the work of Alfriend et al. [5] in assuming that pdf is constant within the integration radius of the combined object,

Klinkrad [13] obtained the maximum collision probability by multiplying the maximum pdf by the area of the combined collision section.

Chan [12] discussed two approaches to determine the maximum probability: the maximum likelihood method and maximum probability method. The first approach can be used to obtain a quick estimate of the maximum probability when the pdf does not vary much over the combined circular collision cross-section area, which is somewhat similar to the process from Alfriend et al. [5]. The second approach is an exact analysis of the problem depending only on the accuracy of the analytical expression for the collision probability. In the second approach the variables are the size and shape. The orientation (represented by the angle between major axes of covariance ellipse and the relative position vector) has been treated as a constant.

Alfano [14] found an “absolute maximum probability” by holding the miss-distance magnitude and the combined object radius fixed and scaling and reorienting the covariance. It was shown that to achieve such a maximum, the relative position vector should lie on the covariance ellipse’s major axis. By specifying the aspect ratio, Alfano [14] forced the shape of covariance ellipse to be fixed. In this sense, the maximum probability obtained by Alfano [14] is not “absolute”.

Laporte [15] presented a maximum probability method that scales each covariance individually to find the maximum probability for fixed shape and orientation of error ellipses. Frisbee [16] investigated a maximum probability approach if the error covariance information for only one of two objects is available. Frisbee [16] used the known covariance information along with a critical value of the missing covariance to obtain an approximate but useful P_c upper bound. The unknown uncertainty is represented by a degenerate ellipse, which is oriented along the relative position vector. Then the critical value of variance that maximizes the probability is determined. The size, shape, and orientation of the known covariance ellipse (always of the primary object) are supposed to be fixed, while those of the unknown covariance are treated as variables to achieve the upper bound of collision probability.

These previous works were insightful from their own aspects, though they did not explicitly clarify what was their assumption on the variation of covariance ellipse. The results gained from these methods may be diverse, which will lead to confusion. In this section, a clear and comprehensive discussion will be given to address all the situations when assessing the maximum collision probability. The discussion will be based on Chan’s analytical equation for calculating the collision probability.

Figure 6.13 categorizes all situations when calculating P_{cmax} according to the changeability of the orientation, shape, and size of covariance ellipse. There are eight situations obtained. Situation No. 1 is not meaningful since it does not yield P_{cmax} because there is no variable factor. Situations No. 3 and 7 are not practical, because the shapes of covariance ellipse are variable but the sizes are fixed. Situations No. 2, 4, 6, and 8 were partially or completely discussed by Alfriend (1999), Chan [12], Alfano [14], and Alfano [9], respectively.

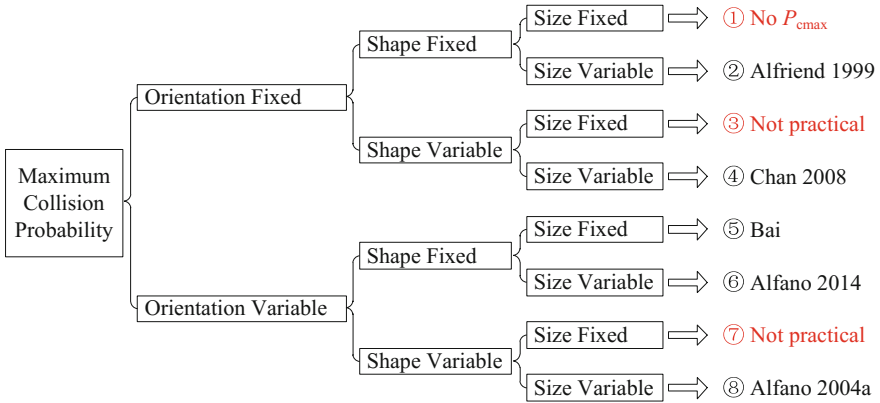


Fig. 6.13 Eight covariance situations for maximum collision probability assessment

This section is organized as follow. First, Chan’s analytical formula of collision probability is given. Then according to Fig. 6.13, situations No. 2, 4, 5, 6, and 8 are discussed with uniform definitions and symbols and independent derivations. The consequences are evaluated by comparison with those obtained from previous works. Finally, a practical conjunction event as a test case is presented to demonstrate and validate the results.

6.2.2 Fixed Orientation

Considering the largest uncertainty is generally in the in-track direction, we can estimate the direction of the principal axis that determines the orientation of error ellipsoid. In this section we assume that the orientation of error ellipse is a constant, but the size and/or shape are variable. P_{cmax} will be analyzed both in the case of fixed error ellipse shape and in the case of variable error ellipse shape.

1. Fixed shape, variable size (Situation No. 2)

The accurate value of the prediction position error of space objects may not be easily gained in practical conjunction risk assessment. However, we usually could know the shape of error ellipse via orbital dynamics and error propagation analysis. Though the accurate value of P_c cannot be ascertained in this case, P_{cmax} can be. In the case of fixed error ellipse shape, aspect ratios of error covariance and the ratio of covariance ellipse’s major axis are predefined.

The explicit expression of collision probability Eqs. (5.22) and (5.23) indicate that the error variances of each object in each direction are simply summed up to form the combined error variances. What really matters in the computation of P_c and P_{cmax} is sum of the error variances of each object in each direction instead of the individual ones. Specific values of combined variances (sum of the error

variances) could lead to certain P_c or $P_{c\max}$, no matter what the individual variance is. Therefore, we could let the sizing of covariances of two objects have the same scaling factor for convenience. If we consider the case in which the covariances of two objects are multiplied by a scaling factor k^2 , and find the value of k which maximizes P_c .

$$\mathbf{P}_{1RSW}(k) = k^2 \begin{bmatrix} \sigma_{1R}^2 & 0 & 0 \\ 0 & \sigma_{1S}^2 & 0 \\ 0 & 0 & \sigma_{1W}^2 \end{bmatrix}, \quad \mathbf{P}_{2RSW}(k) = k^2 \begin{bmatrix} \sigma_{2R}^2 & 0 & 0 \\ 0 & \sigma_{2S}^2 & 0 \\ 0 & 0 & \sigma_{2W}^2 \end{bmatrix} \quad (6.21)$$

where k could be any positive real number.

Substituting Eq. (6.21) into Eq. (5.26), we can obtain

$$P_c(k) = \exp \left[-\frac{1}{2k^2} \left(\frac{R^2}{\sigma_R^2} + \frac{S^2 + W^2}{\sigma_{SW}^2} \right) \right] \cdot \left[1 - \exp \left(-\frac{r_A^2}{2k^2 \sigma_R \sigma_{SW}} \right) \right] \quad (6.22)$$

Hence, the P_c can be expressed as one-variable function of the scaling factor k

$$P_c(k) = e^{-\frac{X^2}{2k^2}} \left(1 - e^{-\frac{Y^2}{2k^2}} \right) \quad (6.23)$$

where

$$X^2 = \frac{R^2}{\sigma_R^2} + \frac{S^2 + W^2}{\sigma_{SW}^2}, \quad Y^2 = \frac{r_A^2}{\sigma_R \sigma_{SW}} \quad (6.24)$$

X and Y can be regarded as dimensionless equivalent closest approach distance and equivalent collision cross-section radius respectively, with non-spherical distribution of position error. There is no unknown parameter in the expression of X and Y .

We take the derivative of $P_c(k)$ with respect to k^2 :

$$\begin{aligned} \frac{\partial P_c(k)}{\partial k^2} &= \frac{X^2}{2k^4} e^{-\frac{X^2}{2k^2}} \left(1 - e^{-\frac{Y^2}{2k^2}} \right) - \frac{Y^2}{2k^4} e^{-\frac{X^2}{2k^2}} e^{-\frac{Y^2}{2k^2}} \\ &= \frac{X^2}{2k^4} e^{-\frac{X^2}{2k^2}} - \frac{X^2 + Y^2}{2k^4} e^{-\frac{X^2 + Y^2}{2k^2}} \end{aligned} \quad (6.25)$$

Set the derivative equals to zero, the value of k^2 which maximizes $P_c(k)$ is the solution to the equations

$$\frac{\partial P_c(k)}{\partial k^2} = \frac{X^2}{2k^4} e^{-\frac{X^2}{2k^2}} - \frac{X^2 + Y^2}{2k^4} e^{-\frac{X^2 + Y^2}{2k^2}} = 0 \quad (6.26)$$

By solving Eq. (6.26) we can get the k^2

$$\begin{aligned} \frac{X^2}{2k^4} e^{-\frac{X^2}{2k^2}} - \frac{X^2 + Y^2}{2k^4} e^{-\frac{X^2 + Y^2}{2k^2}} = 0 &\Leftrightarrow X^2 e^{-\frac{X^2}{2k^2}} = (X^2 + Y^2) e^{-\frac{X^2 + Y^2}{2k^2}} \\ \Leftrightarrow \frac{X^2}{X^2 + Y^2} = e^{-\frac{Y^2}{2k^2}} &\Leftrightarrow \frac{Y^2}{2k^2} = \ln \frac{X^2 + Y^2}{X^2} \Leftrightarrow k^2 = \frac{Y^2}{2 \ln \left(1 + \frac{Y^2}{X^2}\right)} \end{aligned} \quad (6.27)$$

Substituting Eq. (6.27) into Eq. (6.23), the corresponding value of $P_{c\max}$ is

$$P_{c\max} = \left(\frac{X^2}{X^2 + Y^2} \right)^{\frac{X^2}{Y^2}} \left(\frac{Y^2}{X^2 + Y^2} \right) = \left(\frac{\frac{X^2}{Y^2}}{1 + \frac{X^2}{Y^2}} \right)^{\frac{X^2}{Y^2}} \left(\frac{1}{1 + \frac{X^2}{Y^2}} \right) \quad (6.28)$$

The dimensionless parameter λ can be defined as

$$\lambda \equiv X^2/Y^2 \quad (6.29)$$

the $P_{c\max}$ and k^2 could be written as

$$P_{c\max} = \frac{\lambda^\lambda}{(1 + \lambda)^{1 + \lambda}}; \quad k^2 = \frac{X^2}{2\lambda \ln(1 + 1/\lambda)} \quad (6.30)$$

Because combined collision cross-section radius r_A is much less than the closest approach distance in common situations, the inverse of λ , $1/\lambda$, is a small quantity. The limit of k^2 and $P_{c\max}$ when $1/\lambda$ tends to zero are

$$\begin{aligned} k^2 &= \frac{X^2}{2\lambda \ln(1 + 1/\lambda)} \rightarrow \frac{X^2}{2} = \frac{1}{2} \left(\frac{R^2}{\sigma_R^2} + \frac{S^2 + W^2}{\sigma_{SW}^2} \right), \quad \lambda \rightarrow \infty \\ \lambda P_{c\max} &= \left(\frac{\lambda}{1 + \lambda} \right)^{1 + \lambda} = \left(1 - \frac{1}{1 + \lambda} \right)^{(1 + \lambda)} \rightarrow \frac{1}{e}, \quad P_{c\max} \rightarrow \frac{1}{e\lambda}, \quad \lambda \rightarrow \infty \end{aligned} \quad (6.31)$$

The relative error of the approximate $P_{c\max}$ in Eq. (6.31) to $P_{c\max}$ in Eq. (6.30) is about 9.85% when $\lambda = 5$, and 4.96% when $\lambda = 10$. The larger λ is, the smaller the relative error will be. The miss distance of a conjunction is generally of the magnitude of kilometers, while the combined collision cross-section radius r_A is less than 100 m. Therefore the relative error of the approximate expression is commonly less than 5%. This accuracy reveals that the approximate expression of $P_{c\max}$ is useful in preliminary analysis.

When the error scaling factor k^2 that maximizes P_c is known, it is easy to determine if there is a probability dilution based on the relationship between k and 1. P_c is plotted as a function of the scaling factor in Fig. 6.14. When scaling factor k increases P_c first gradually increase to a maximum value $P_{c\max}$, and then decrease. It is obvious that P_c is not diluted before approaching $P_{c\max}$ and it is more sensitive than the right.

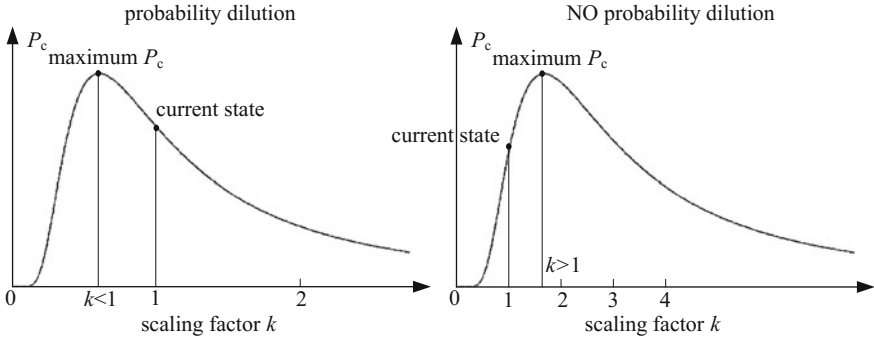


Fig. 6.14 Collision probability dilution assessment

The state $k = 1$ represents current error covariance state. If $P_{c\max}$ corresponded $k < 1$ (as seen in Fig. 6.14's left side), it indicates that current error covariance is larger than the error covariance that yield $P_{c\max}$, this is a probability dilution situation. P_c will decrease as the error covariance increases. If $P_{c\max}$ corresponded $k > 1$ (as seen in the right side of Fig. 6.14), on the other hand, it indicates that current error covariance is smaller than the error covariance that yield $P_{c\max}$, there is no probability dilution. The P_c will increase as the error covariance increases.

Alfriend et al. [5] derived the maximized P_c and corresponding value of scaling factor k^2 using an approximate collision probability calculated by assuming the probability density is constant over the collision circle.

$$k^2 = \frac{s_o^T P^{*-1} s_o}{2}, \quad P_{c\max} = \frac{r_A^2}{e(s_o^T P^{*-1} s_o) |P^*|^{1/2}} \quad (6.32)$$

where s_o is the estimated relative position vector in the conjunction plane, and P^* is the 2×2 covariance matrix in the encounter system. By considering the following form of s_o and P^*

$$s_o = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \quad P^* = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \quad (6.33)$$

Equation (6.32) can be rewritten as

$$k^2 = \frac{1}{2} \left(\frac{\mu_x^2}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2} \right), \quad P_{c\max} = \frac{\frac{r_A^2}{\sigma_x \sigma_y}}{e \left(\frac{\mu_x^2}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2} \right)} \quad (6.34)$$

Considering Eq. (6.24), the scaling factor k^2 and corresponding $P_{c\max}$ in Eq. (6.34) can be rewritten as

$$k^2 = \frac{X^2}{2}, \quad P_{c \max} = \frac{Y^2}{eX^2} = \frac{1}{e\lambda} \quad (6.35)$$

Comparing Eqs. (6.31) with (6.35), it is obvious that the approximate expression of $P_{c \max}$ and scaling factor k^2 are identical to Alfriend's results in the limit. Therefore, the constant probability density approximation method is an extreme case of Chan's method when the combined collision cross-section radius r_A is much less than the closest approach distance. As discussed before, the relative error of Alfriend et al.'s [5] results are less than 5% for most conjunctions, but the general expression of $P_{c \max}$ and scaling factor k^2 in Eq. (6.31) are more accurate when the combined radius is close to closest approach distance.

2. Variable shape, variable size (Situation No. 4)

In this case, not only the size but also the shape of error ellipse can be varied. The calculation method for $P_{c \max}$ and corresponding error standard deviation with fixed error ellipse's shape is the maximization of a one-variable function. On the other hand, the calculation method for variable error ellipse shape is the maximization of a multivariable function. From Eq. (5.9) P_c can be expressed as a function of two independent combined error standard deviations, σ_R and σ_{SW} , which are in the radial direction and in the horizontal plane, respectively. The function $P_c = P_c(\sigma_R, \sigma_{SW})$ depends on two independent variables. The maximum probability and corresponding combined error standard deviations can be obtained by maximizing the two-variable function

$$P_c(\sigma_x, \sigma_y) = \exp \left[-\frac{1}{2} \left(\frac{\mu_x^2}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2} \right) \right] \left[1 - \exp \left(-\frac{r_A^2}{2\sigma_x\sigma_y} \right) \right] \quad (6.36)$$

Next step is to determine the standard deviations σ_x and σ_y that maximize the collision probability function $P_c(\sigma_x, \sigma_y)$. Similar to the one-variable case discussed above, the partial derivatives of $P_c = P_c(\sigma_x, \sigma_y)$ with respect to σ_x and σ_y are expressed as following:

$$\begin{cases} \frac{\partial P_c}{\partial \sigma_x} = e^{-\frac{1}{2} \left(\frac{\mu_x^2}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2} \right)} \left(1 - e^{-\frac{r_A^2}{2\sigma_x\sigma_y}} \right) \frac{\mu_x^2}{\sigma_x^3} - e^{-\frac{1}{2} \left(\frac{\mu_x^2}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2} \right)} e^{-\frac{r_A^2}{2\sigma_x\sigma_y}} \frac{r_A^2}{2\sigma_x\sigma_y} \\ \frac{\partial P_c}{\partial \sigma_y} = e^{-\frac{1}{2} \left(\frac{\mu_x^2}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2} \right)} \left(1 - e^{-\frac{r_A^2}{2\sigma_x\sigma_y}} \right) \frac{\mu_y^2}{\sigma_y^3} - e^{-\frac{1}{2} \left(\frac{\mu_x^2}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2} \right)} e^{-\frac{r_A^2}{2\sigma_x\sigma_y}} \frac{r_A^2}{2\sigma_x\sigma_y} \end{cases} \quad (6.37)$$

The necessary condition of $P_{c \max}$ is that these two partial derivatives equal zero simultaneously.

$$\frac{\partial P_c}{\partial \sigma_x} = \frac{\partial P_c}{\partial \sigma_y} = 0 \quad (6.38)$$

To obtain the values of σ_x and σ_y , the two nonlinear equations Eq. (6.38) have to be solved. Substituting Eq. (6.37) into Eq. (6.38), through simplifying we can get

$$\begin{cases} \frac{\mu_x^2}{\sigma_x} e^{\frac{r_A^2}{2\sigma_x\sigma_y}} = \frac{\mu_x^2}{\sigma_x} + \frac{r_A^2}{2\sigma_y} \\ \frac{\mu_y^2}{\sigma_y} e^{\frac{r_A^2}{2\sigma_x\sigma_y}} = \frac{\mu_y^2}{\sigma_y} + \frac{r_A^2}{2\sigma_x} \end{cases} \quad (6.39)$$

From the second equation in Eq. (6.39) we know

$$e^{\frac{r_A^2}{2\sigma_x\sigma_y}} = 1 + \frac{r_A^2\sigma_y}{2\mu_y^2\sigma_x} \quad (6.40)$$

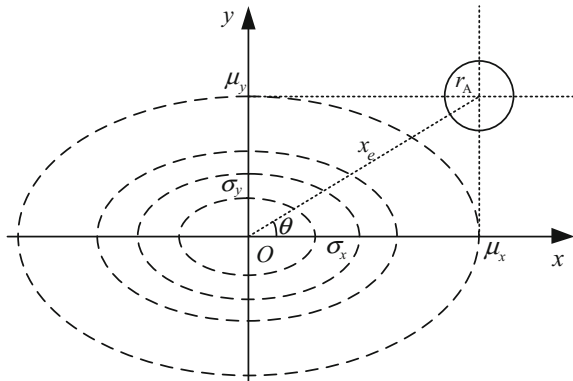
Substituting Eq. (6.40) into the first equation in Eq. (6.39), we obtain the simple relationship between σ_x and σ_y

$$\frac{\mu_x^2}{\mu_y^2} = \frac{\sigma_x^2}{\sigma_y^2}, \quad \sigma_y = \frac{|\mu_y|}{|\mu_x|} \sigma_x \quad (6.41)$$

Equation (6.41) demonstrates that the values of σ_x and σ_y lie on the line which cross the point of closest approach in the conjunction plane. As seen in Fig. 6.15, the shape of the error ellipse which maximizes P_c is similar to the largest dash ellipse with the major and minor axes μ_x and μ_y . Substitution of Eq. (6.41) into the first equation in Eq. (6.39) yields

$$\begin{cases} \frac{1}{\sigma_x^2} = \frac{2|\mu_y|}{r_A^2|\mu_x|} \ln\left(1 + \frac{r_A^2}{2|\mu_x||\mu_y|}\right) \\ \frac{1}{\sigma_y^2} = \frac{2|\mu_x|}{r_A^2|\mu_y|} \ln\left(1 + \frac{r_A^2}{2|\mu_x||\mu_y|}\right) \end{cases} \quad (6.42)$$

Fig. 6.15 Parameters of P_c density function and radius of collision cross-section area in the conjunction plane



Equations (6.42) give the closed-form expressions for σ_x and σ_y in terms of the collision parameters μ_x , μ_y and r_A . Substitution of Eqs. (6.42) into Eq. (6.7) yields

$$P_{c \max} = \left(1 + \frac{r_A^2}{2|\mu_x||\mu_y|}\right)^{-\frac{2|\mu_x||\mu_y|}{r_A^2}} \left(\frac{\frac{r_A^2}{2|\mu_x||\mu_y|}}{1 + \frac{r_A^2}{2|\mu_x||\mu_y|}}\right) \quad (6.43)$$

Define the dimensionless parameter η as

$$\eta \equiv \frac{2|\mu_x||\mu_y|}{r_A^2} \quad (6.44)$$

By substituting Eq. (6.44) into Eq. (6.43), we obtain the compact expression of $P_{c \max}$

$$P_{c \max} = \frac{\eta^\eta}{(1 + \eta)^{1 + \eta}} \quad (6.45)$$

The corresponding σ_x and σ_y are

$$\sigma_x^2 = \frac{\mu_x^2}{\eta \ln(1 + 1/\eta)}, \quad \sigma_y^2 = \frac{\mu_y^2}{\eta \ln(1 + 1/\eta)} \quad (6.46)$$

By replacing μ_x and μ_y with the components of relative positions in the RSW coordinate system at the TCA, i.e., substituting Eq. (5.26) into Eqs. (6.45) and (6.46), we obtain the new expressions of dimensionless parameter η and combined error standard deviations σ_R and σ_{SW} that maximize P_c .

$$\eta = \frac{2|R|\sqrt{S^2 + W^2}}{r_A^2}, \quad \sigma_R = \frac{|R|}{\sqrt{\eta \ln(1 + 1/\eta)}}, \quad \sigma_{SW} = \frac{\sqrt{S^2 + W^2}}{\sqrt{\eta \ln(1 + 1/\eta)}} \quad (6.47)$$

Note that Eq. (6.45) only involves the dimensionless parameter η . Equations (6.45) and (6.47) are used to compute $P_{c \max}$ and corresponding combined error standard deviations when R , S , W , and r_A are specified.

Because combined r_A is smaller than the closest approach distance in almost all cases, the inverse of η , $1/\eta$, is a small quantity. When $1/\eta$ tends to zero, we may write the limits of $P_{c \max}$, σ_x and σ_y as

$$\left. \begin{aligned} P_{c \max} &= \frac{\eta^\eta}{(1 + \eta)^{1 + \eta}} \rightarrow \frac{1}{e\eta} = \frac{r_A^2}{2e|\mu_x||\mu_y|} \\ \sigma_x &= \frac{|\mu_x|}{\sqrt{\eta \ln(1 + 1/\eta)}} \rightarrow |\mu_x| \\ \sigma_y &= \frac{|\mu_y|}{\sqrt{\eta \ln(1 + 1/\eta)}} \rightarrow |\mu_y| \end{aligned} \right\} \eta \rightarrow \infty \quad (6.48)$$

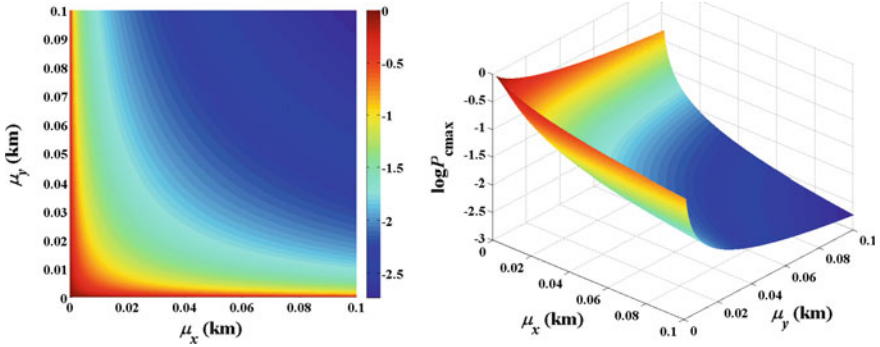


Fig. 6.16 Two-dimensional distribution (*left*) and three-dimensional mesh surface (*right*) of denary logarithm of P_{cmax} with respect to μ_x and μ_y when $r_A = 10$ m

These approximate expressions of combined error standard deviations are useful in preliminary analysis of position error yield P_{cmax} .

Comparing Eqs. (6.30) and (6.45), we note that the expressions of P_{cmax} in the case of fixed error ellipse shape and in the case of variable error ellipse shape are uniform, only the definitions of dimensionless parameters λ and η are different. $P_{cmax}(\eta)$ is a decreasing function with respect to η . Figure 6.16 illustrates the two-dimensional distribution (left) and three-dimensional mesh surface (right) of denary logarithm of P_{cmax} with respect to μ_x and μ_y when $r_A = 10$ m.

Moreover, Eq. (6.48) indicates that P_{cmax} is approximately in direct proportion to the square of collision cross-section radius, r_A^2 , and in inverse proportion to $|R|\sqrt{S^2 + W^2}$, which has the same dimension as the square of the closest approach distance, and is of the same order of magnitude. Let $r_A = 20$ m, the P_{cmax} is calculated by Eq. (6.45) at different relative position (radial, in-track, cross-track) at TCA. The results are summarized in Table 6.10, which corresponds well with the inverse proportion property of the P_{cmax} to relative position.

Chan [12] obtained the closed-form expression of maximum collision probability and corresponding standard deviations in terms of collision parameters r_A , x_e , and θ .

Table 6.10 The P_{cmax} calculated using different relative position

Radial (R)/km	In-track (S)/km	Cross-track (W)/km	P_{cmax}
0.1	3	0.3	2.4395e-04
1	30	3	2.4404e-06
2	60	6	6.1009e-07
5	150	15	9.7614e-08
10	300	30	2.4404e-08
20	600	60	6.1009e-09
50	1500	150	9.7614e-10
100	3000	300	2.4404e-10

$$\begin{aligned}
 P_{c \max}(r_A, x_e, \theta) &= \left(\frac{r_A^2}{r_A^2 + x_e^2 \sin^2 2\theta} \right) \left(\frac{x_e^2 \sin 2\theta}{r_A^2 + x_e^2 \sin^2 2\theta} \right)^{\frac{x_e^2(1-\cos 2\theta)}{r_A^2 \tan \theta}} \\
 \sigma_x^2 &= \frac{r_A^2}{2 \tan \theta \ln \left(1 + \frac{r_A^2}{x_e^2 \sin^2 2\theta} \right)} \\
 \sigma_y^2 &= \frac{r_A^2 \tan \theta}{2 \ln \left(1 + \frac{r_A^2}{x_e^2 \sin^2 2\theta} \right)}
 \end{aligned} \tag{6.49}$$

where x_e is the distance of the closest approach or the miss distance, and θ the angle between the point of closest approach and the principal axis of error ellipse as illustrated in Fig. 6.15.

Chan's collision parameters x_e and θ and the probability integral parameters μ_x and μ_y in this section have the relationship as follow

$$\mu_x = x_e \cos \theta, \quad \mu_y = x_e \sin \theta \tag{6.50}$$

Substitution of Eq. (6.50) into Eq. (6.49) yields

$$\begin{aligned}
 P_{c \max}(r_A, x_e, \theta) &= \left(\frac{r_A^2}{r_A^2 + 2|\mu_x||\mu_y|} \right) \left(\frac{2|\mu_x||\mu_y|}{r_A^2 + 2|\mu_x||\mu_y|} \right)^{\frac{2|\mu_x||\mu_y|}{r_A^2}} = \frac{\eta^\eta}{(1+\eta)^{1+\eta}} \\
 \sigma_x^2 &= \frac{r_A^2 \mu_x^2}{2|\mu_x||\mu_y| \ln \left(1 + \frac{r_A^2}{2|\mu_x||\mu_y|} \right)} = \frac{\mu_x^2}{\eta \ln(1+1/\eta)} \\
 \sigma_y^2 &= \frac{r_A^2 \mu_y^2}{2|\mu_x||\mu_y| \ln \left(1 + \frac{r_A^2}{2|\mu_x||\mu_y|} \right)} = \frac{\mu_y^2}{\eta \ln(1+1/\eta)}
 \end{aligned} \tag{6.51}$$

In Chan [12]'s maximum likelihood approach, the pdf does not vary much over the combined circular collision cross-section, $P_{c \max}$ and the standard deviations that maximize the collision probability are approximately given by

$$P_{c \max} = \frac{r_A^2}{e x_e^2 \sin^2 2\theta}, \quad \sigma_x = x_e |\cos \theta|, \quad \sigma_y = x_e |\sin \theta| \tag{6.52}$$

substitution of Eqs. (6.50)–(6.52) yields the approximate $P_{c \max}$ and the corresponding standard deviations as

$$P_{c \max} = \frac{r_A^2}{2e|\mu_x||\mu_y|} = \frac{1}{e\eta}, \quad \sigma_x = |\mu_x|, \quad \sigma_y = |\mu_y| \tag{6.53}$$

In Chan [12]’s work the three-dimensional ellipsoids of two objects can be of arbitrary orientation, shape, and size. The only constraint is that the relative position in the conjunction plane is fixed because the nominal positions of two objects are very well known from their respective orbit determinations. The rest of the analysis is general. That is, the covariance ellipsoids of the two conjuncting objects can have arbitrary orientations, shapes, and sizes.

Equation (6.51) indicates that the results of Chan [12] and those derived in this section are essentially identical. However, the explicit expressions of P_{cmax} and corresponding error standard deviations presented in this section still have their significance. First, this section expresses P_c and P_{cmax} in terms of components of relative position in the RSW frame instead of integral parameters in the conjunction plane, and directly connects the two criteria (miss distance or its components and P_c) for conjunction risk assessment. It is beneficial to an operator because the connection should be considered in P_c sensitivity analysis, the establishment of relationship between distance threshold and probability threshold, and the choice of a proper distance threshold. Second, the introduction of dimensionless parameters λ and η unifies the form of P_{cmax} both in the fixed error ellipse shape case and in the variable error ellipse shape case, and simplifies P_{cmax} and corresponding error standard deviations into a more compact and understandable expression.

6.2.3 Variable Orientation

When the size, shape, and orientation of covariance ellipse are variable, the certain factors are the miss distance x_e and the radius r_A . By substituting Eq. (6.50) into Eq. (5.9), the expression of collision probability can be written as

$$P_c = \exp \left[-\frac{x_e^2}{2} \left(\frac{\cos^2 \theta}{\sigma_x^2} + \frac{\sin^2 \theta}{\sigma_y^2} \right) \right] \left[1 - \exp \left(-\frac{r_A^2}{2\sigma_x \sigma_y} \right) \right] \quad (6.54)$$

the first and second-order partial derivatives of P_c with respect to θ are

$$\begin{aligned} \frac{\partial P_c}{\partial \theta} &= P_c \frac{x_e^2 \sin 2\theta}{2} \left(\frac{1}{\sigma_x^2} - \frac{1}{\sigma_y^2} \right) \\ \frac{\partial^2 P_c}{\partial \theta^2} &= P_c x_e^2 \cos 2\theta \left(\frac{1}{\sigma_x^2} - \frac{1}{\sigma_y^2} \right) \end{aligned} \quad (6.55)$$

When $\theta = 0$ or $\theta = \pi/2$, the partial derivative $\partial P_c / \partial \theta = 0$, where θ is the angle between the point of closest approach and the x -axis as illustrated in Fig. 5.2. If $\sigma_x > \sigma_y$, the x -axis is the major axis of the error ellipse, then the θ maximizing P_c is $\theta = 0$ ($\partial P_c / \partial \theta = 0$, $\partial^2 P_c / \partial \theta^2 < 0$); If $\sigma_x < \sigma_y$, the x -axis is the minor axis of the

Table 6.11 The maximum and minimum P_c with respect to the angle θ

	$\sigma_x > \sigma_y$	$\sigma_x < \sigma_y$
$\theta = 0$	$\frac{\partial^2 P_c}{\partial \theta^2} < 0, P_{cmax}$	$\frac{\partial^2 P_c}{\partial \theta^2} > 0, P_{cmin}$
$\theta = \pi/2$	$\frac{\partial^2 P_c}{\partial \theta^2} > 0, P_{cmin}$	$\frac{\partial^2 P_c}{\partial \theta^2} < 0, P_{cmax}$

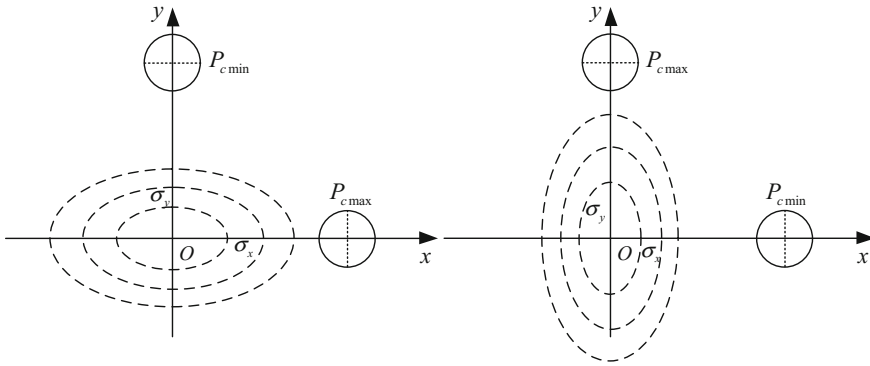


Fig. 6.17 The maximum and minimum P_c with respect to the angle θ

error ellipse, then the θ maximizing P_c is $\theta = \pi/2$ ($\partial P_c / \partial \theta = 0, \partial^2 P_c / \partial \theta^2 < 0$). Table 6.11 shows four cases about the maximum and the minimum P_c with respect to the angle θ .

Table 6.11 shows that the P_c maximized when the center of collision cross-section area is on the major axis of error ellipse, while the P_c minimized when the center of collision cross-section area is on the minor axis of error ellipse; as illustrated in Fig. 6.17.

1. Fixed shape, fixed size (Situation No. 5)

When the center of collision cross-section area is on the major axis of error ellipse ($\theta = 0, \sigma_x > \sigma_y$, or $\theta = \pi/2, \sigma_x < \sigma_y$), the expression of maximum P_c is

$$P_{cmax} = \exp\left(-\frac{x_c^2}{2 \max(\sigma_x^2, \sigma_y^2)}\right) \left[1 - \exp\left(-\frac{r_A^2}{2\sigma_x\sigma_y}\right)\right] \tag{6.56}$$

Equation (6.56) is the expression of P_{cmax} in the case of fixed size and fixed shape but variable orientation of covariance ellipse.

2. Fixed shape, variable size (Situation No. 6)

We are usually able to obtain the shape of error ellipse via orbital dynamics and error propagation analysis. In the case of fixed error ellipse shape, the aspect ratio of error covariance is predefined. By convention, let $\sigma_x \geq \sigma_y$, then define the aspect ratio α as

$$\alpha = \frac{\sigma_x}{\sigma_y}, \quad \alpha \geq 1 \quad (6.57)$$

Substituting Eq. (6.57) into Eq. (6.56), we can obtain the expression of $P_{c\max}$ as

$$P_{c\max}(\sigma_x^2) = \exp\left(-\frac{x_e^2}{2\sigma_x^2}\right) \left[1 - \exp\left(-\frac{\alpha r_A^2}{2\sigma_x^2}\right)\right] \quad (6.58)$$

For the case of fixed shape variable size, the aspect ratio α as well as the magnitude of the miss distance x_e are constants, only the error variance σ_x^2 which represent the size of error ellipse is variable.

The derivative of $P_{c\max}(\sigma_x^2)$ with respect to σ_x^2 is:

$$\frac{dP_{c\max}}{d\sigma_x^2} = \frac{x_e^2}{2\sigma_x^4} e^{-\frac{x_e^2}{2\sigma_x^2}} - \frac{x_e^2 + \alpha r_A^2}{2\sigma_x^4} e^{-\frac{x_e^2 + \alpha r_A^2}{2\sigma_x^2}} \quad (6.59)$$

Set the derivative equals to zero, the value of σ_x^2 which maximizes $P_c(\sigma_x^2)$ is the solution to the equations

$$\frac{dP_{c\max}}{d\sigma_x^2} = \frac{x_e^2}{2\sigma_x^4} e^{-\frac{x_e^2}{2\sigma_x^2}} - \frac{x_e^2 + \alpha r_A^2}{2\sigma_x^4} e^{-\frac{x_e^2 + \alpha r_A^2}{2\sigma_x^2}} = 0 \quad (6.60)$$

Solving Eq. (6.60) the σ_x^2 is

$$\sigma_x^2 = \frac{\alpha r_A^2}{2 \ln\left(1 + \frac{\alpha r_A^2}{x_e^2}\right)} \quad (6.61)$$

Substituting Eq. (6.61) into Eq. (6.58), the corresponding value of $P_{c\max}$ is

$$P_{c\max}(\sigma_x^2) = \left(1 + \frac{\alpha r_A^2}{x_e^2}\right)^{-\frac{x_e^2}{\alpha r_A^2}} \left[\frac{\alpha r_A^2}{x_e^2 + \alpha r_A^2}\right] \quad (6.62)$$

Define the dimensionless parameter β as

$$\beta = \frac{\alpha r_A^2}{x_e^2} = \frac{r_A^2 \sigma_x}{x_e^2 \sigma_y} \quad (6.63)$$

the $P_{c\max}$ and corresponding error variances could be written as

$$P_{c \max} = \left(\frac{\beta}{1 + \beta} \right) \left(\frac{1}{1 + \beta} \right)^{\frac{1}{\beta}} \tag{6.64}$$

$$\sigma_x^2 = \frac{\alpha r_A^2}{2 \ln(1 + \beta)}, \quad \sigma_y^2 = \frac{r_A^2}{2\alpha \ln(1 + \beta)}$$

Equations (6.63) and (6.64) are singular when $x_e = 0$, the $P_{c \max}$ equals one in this rare case.

Alfano [14] found maximum probability by holding the miss-distance magnitude and combined object radius fixed and scaling and reorienting the covariance. It was shown that to achieve such a maximum, the miss-distance vector should lie on the covariance ellipse’s major axis. The results given by Alfano [14] are

$$P_{c \max} = \left(\frac{\beta}{1 + \beta} \right) \left(\frac{1}{1 + \beta} \right)^{\frac{1}{\beta}} \tag{6.65}$$

$$\sigma_y = \sqrt{\frac{-r_A^2}{2 \cdot AR \cdot \ln \frac{d^2}{d^2 + AR \cdot r_A^2}}} = \sqrt{\frac{r_A^2}{2\alpha \ln(1 + \beta)}}$$

Equation (6.64) exactly matches the Eq. (6.65) in Alfano’s [14] results. Essentially, Alfano’s [14] maximum probability was found in the case of variable orientation, fixed shape, and variable size of error ellipse.

3. Variable shape, variable size (Situation No. 8)

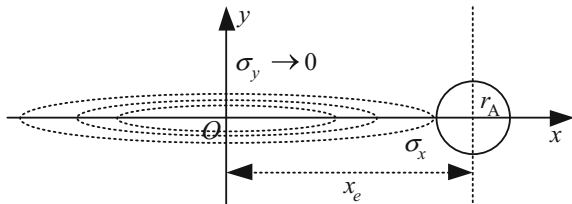
For the case of variable orientation, variable shape, and variable size, the constants are the miss distance x_e and the radius of collision section area. In Eq. (6.58) both the error variance σ_x^2 which represent the size of error ellipse and the aspect ratio α which represent the shape of error ellipse are variable.

We take the derivative of $P_{c \max}(\alpha)$ with respect to α :

$$\frac{dP_{c \max}}{d\alpha} = \exp\left(-\frac{x_e^2 + \alpha r_A^2}{2\sigma_x^2}\right) \frac{r_A^2}{2\sigma_x^2} > 0 \tag{6.66}$$

The derivative in Eq. (6.66) is always positive, hence $P_{c \max}(\alpha)$ is an increasing function with respect to aspect ratio α . The limiting case is that $\sigma_y \rightarrow 0$ therefore $\alpha \rightarrow \infty$. For this case the problem has been reduced from 2-dimensional to 1-dimensional, as illustrated in Fig. 6.18.

Fig. 6.18 The limit case when $\sigma_y \rightarrow 0$



For the 1-dimensional case, Eq. (5.1) can be rewritten as

$$P_c = \int_{x_e - r_A}^{x_e + r_A} \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{x^2}{2\sigma_x^2}} dx = \int_{\frac{x_e - r_A}{\sigma_x}}^{\frac{x_e + r_A}{\sigma_x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \Phi\left(\frac{x_e + r_A}{\sigma_x}\right) - \Phi\left(\frac{x_e - r_A}{\sigma_x}\right) \quad (6.67)$$

Let the derivatives of $P_c(\sigma_x)$ equals to zero

$$\frac{dP_c(\sigma_x)}{d\sigma_x} = -\varphi\left(\frac{x_e + r_A}{\sigma_x}\right) \frac{x_e + r_A}{\sigma_x^2} + \varphi\left(\frac{x_e - r_A}{\sigma_x}\right) \frac{x_e - r_A}{\sigma_x^2} = 0 \quad (6.68)$$

So

$$\sigma_x^2 = \frac{2x_e r_A}{\ln \frac{x_e + r_A}{x_e - r_A}} \quad (6.69)$$

Let the dimensionless parameter $\delta = r_A/x_e$, Eq. (6.69) becomes

$$\sigma_x^2 = x_e^2 \frac{2\delta}{\ln \frac{1+\delta}{1-\delta}}, \quad \hat{\sigma} = \frac{\sigma_x}{x_e} = \sqrt{\frac{2\delta}{\ln \frac{1+\delta}{1-\delta}}} \quad (6.70)$$

where $\hat{\sigma}$ is the normalized error standard deviation by miss distance. Substitution of Eq. (6.70) into Eq. (6.67) yields

$$P_{c \max} = \Phi\left(\frac{1+\delta}{\sqrt{2\delta}} \sqrt{\ln \frac{1+\delta}{1-\delta}}\right) - \Phi\left(\frac{1-\delta}{\sqrt{2\delta}} \sqrt{\ln \frac{1+\delta}{1-\delta}}\right) \quad (6.71)$$

The relationship between the standard Gaussian distribution function $\Phi(x)$ and the error function $\text{erf}(x)$ is

$$\Phi(x) = \frac{\text{erf}\left(\frac{x}{\sqrt{2}}\right) + 1}{2} \quad (6.72)$$

Then Eq. (6.71) can be written using error function as

$$P_{c \max} = \frac{1}{2} \text{erf}\left(\frac{\delta+1}{2\sqrt{\delta}} \sqrt{\ln \frac{1+\delta}{1-\delta}}\right) + \frac{1}{2} \text{erf}\left(\frac{\delta-1}{2\sqrt{\delta}} \sqrt{\ln \frac{1+\delta}{1-\delta}}\right) \quad (6.73)$$

In Chan [12]'s maximum likelihood approach, $P_{c \max}$ was approximately given by Eq. (6.52). The right-hand side of Eq. (6.52) is undefined when $\theta = 0$, thus

necessitating the use of the Dirac delta function. In this case, Eq. (6.52) can be replaced by

$$P_{c \max} = \sqrt{\frac{2}{\pi e}} \left(\frac{r_A}{x_e} \right) \approx 0.48394\delta \quad (6.74)$$

Equation (6.74) is exactly the limit of Eqs. (6.71) and (6.73) when $\delta \rightarrow 0$.

Alfano [9] has derived $P_{c \max}$ for the case of $\theta = 0$. Alfano [9] showed how to calculate the upper bounds of probability by determining the “worst” possible covariance parameters and orientation. He assumed the worst upper bound of P_c by taking the one-dimensional pdf on the x -axis so that there is no two-dimensional ellipse in the conjunction plane. The $P_{c \max}$ given by Alfano [9] can be regarded as the “absolute” maximum probability. However, in most cases of close approach, the angle 9θ is not close to zero. In general, the upper bound of P_c is often two to four orders of magnitude larger than those in the relevant range $10^\circ \leq |\theta| \leq 90^\circ$. The upper bound of P_c in Alfano [9]’s extreme situation is often overestimated, which may be too conservative to serve as the discriminators for pre-filtering.

In summarizing, Table 6.12 shows the formulae of $P_{c \max}$ and corresponding error covariances (scaling or orientating factor) of the five situations discussed above.

6.2.4 Examples

To demonstrate the calculation method of $P_{c \max}$, this section will take the U.S. and Russian satellites’ collision event for example. The parameters and details of the event are shown in Sect. 5.2.4. In the condition represented by Tables 5.1, 5.2 and 5.3, P_c between two objects was calculated by means of Eq. (5.9) as $P_c = 1.807975 \times 10^{-4}$.

1. Situation no. 2—fixed orientation, fixed shape, variable size

In this case we calculate $P_{c \max}$ and the corresponding error scaling factor by Eq. (6.30). The resulted $P_{c \max}$ and corresponding error scaling factor and error standard deviations of two objects are shown in Table 6.13.

$P_{c \max}$ corresponded $k = 1.756 > 1$ (as shown in Fig. 6.19), it indicates that the current error covariance ($k = 1$) is smaller than the error covariance that yield $P_{c \max}$. Therefore the current state is not in the probability dilution region. P_c increases as the error covariance increases.

2. Situation no. 4—fixed orientation, variable shape, variable size

In this case we calculate the $P_{c \max}$ and the corresponding combined error standard deviations by Eqs. (6.45) and (6.46). The resulted $P_{c \max}$ and corresponding combined error standard deviations are shown in Table 6.14.

Table 6.12 Summary of the results

Situation	Situation description	Dimensionless parameter	Maximum probability	Corresponding error covariance (scaling or orientating factor)
1	Fixed orientation	-	$P_c = \exp \left[-\frac{1}{2} \left(\frac{\mu_x^2}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2} \right) \right] \left[1 - \exp \left(-\frac{r_A^2}{2\sigma_x\sigma_y} \right) \right]$	-
	Fixed shape			
	Fixed size			
2	Fixed orientation	$\lambda \equiv \frac{\frac{\mu_x^2 + \mu_y^2}{\sigma_x^2 \sigma_y^2}}{\frac{r_A^2}{\sigma_x \sigma_y}}$	$P_{c \max} = \frac{\lambda^\lambda}{(1+\lambda)^{1+\lambda}}$	$k^2 = \frac{\frac{\mu_x^2}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2}}{2\lambda \ln(1+1/\lambda)}$
	Fixed shape			
	Variable size			
4	Fixed orientation	$\eta \equiv \frac{2 \mu_x \mu_y }{r_A^2}$	$P_{c \max} = \frac{\eta^\eta}{(1+\eta)^{1+\eta}}$	$\sigma_x^2 = \frac{\mu_x^2}{\eta \ln(1+1/\eta)}$ $\sigma_y^2 = \frac{\mu_y^2}{\eta \ln(1+1/\eta)}$
	Variable shape			
	Variable size			
5	Variable orientation	-	$P_{c \max} = \exp \left(-\frac{x_e^2}{2 \max(\sigma_x^2, \sigma_y^2)} \right) \left[1 - \exp \left(-\frac{r_A^2}{2\sigma_x\sigma_y} \right) \right]$	$\theta = 0 (\sigma_x > \sigma_y)$ $\theta = \pi/2 (\sigma_x < \sigma_y)$
	Fixed shape			
	Fixed size			
6	Variable orientation	$\alpha = \frac{\sigma_x}{\sigma_y}, \quad \alpha \geq 1$ $\beta = \frac{\alpha r_A^2}{x_e^2}$	$P_{c \max} = \left(\frac{\beta}{1+\beta} \right) \left(\frac{1}{1+\beta} \right)^{\frac{1}{\beta}}$	$\sigma_x^2 = \frac{\alpha r_A^2}{2 \ln(1+\beta)}$ $\sigma_y^2 = \frac{r_A^2}{2\alpha \ln(1+\beta)}$
	Fixed shape			
	Variable size			
8	Variable orientation	$\delta = \frac{r_A}{x_e} = \sqrt{\alpha^2 + \alpha^{\frac{2}{\beta}}}$	$P_{c \max} = \frac{1}{2} \operatorname{erf} \left(\frac{\delta+1}{2\sqrt{\delta}} \sqrt{\ln \frac{1+\delta}{1-\delta}} \right) + \frac{1}{2} \operatorname{erf} \left(\frac{\delta-1}{2\sqrt{\delta}} \sqrt{\ln \frac{1+\delta}{1-\delta}} \right)$	$\sigma_x^2 = x_e^2 \frac{2\delta}{\ln \frac{1+\delta}{1-\delta}}$ $\sigma_y^2 \rightarrow 0$
	Variable shape			
	Variable size			

Table 6.13 P_{cmax} and corresponding error scaling factor and standard deviations (Situation no. 2)

Object	σ_R/km	σ_S/km	σ_W/km
Primary	0.040600	0.362071	0.126394
Secondary	0.063785	0.720332	0.059904
k	1.756027		
P_{cmax}	4.710037×10^{-4}		

Fig. 6.19 The P_{cmax} and corresponding k indicate that the current state is not in the probability dilution region

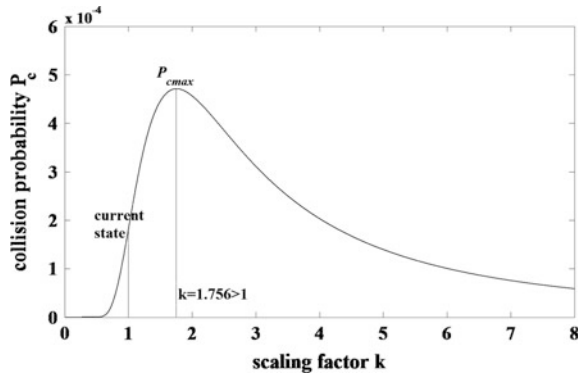


Table 6.14 The P_{cmax} and corresponding combined error standard deviations (Situation no. 4)

P_{cmax}	8.303965×10^{-4}
σ_R/km	0.031748
σ_{SW}/km	0.697688

To demonstrate P_{cmax} and the corresponding combined error standard deviations, the relationship between P_c and the combined error standard deviations σ_R and σ_{SW} is graphically expressed using explicit expression Eq. (5.9). Figure 6.20 illustrates the two-dimensional P_c distribution, and Fig. 6.21 illustrates three-dimensional P_c mesh surface with respect to combined error standard deviations σ_R and σ_{SW} observed from different direction. The empirical P_{cmax} and corresponding combined error standard deviations obtained from Figs. 6.20 and 6.21 are 8.304292×10^{-4} and $\sigma_R = 0.031750 \text{ km}$, $\sigma_{SW} = 0.697500 \text{ km}$. The relative error of P_{cmax} is 0.0039%.

From Table 5.2 $R = 0.031730 \text{ km}$ and $\sqrt{S^2 + W^2} = 0.697294 \text{ km}$. It indicates that the approximate expressions of combined error standard deviations in Eq. (6.48) are correct. We can also calculate the approximate P_{cmax} by approximate expressions in Eq. (6.48), the result $P_{cmax} = 8.313674 \times 10^{-4}$ agrees well with the result in Eq. (6.45). The relative error of approximate P_{cmax} is 0.1%.

Table 6.15 compares P_{cmax} and the corresponding combined error standard deviations results obtained from the empirical mesh surface in Figs. 6.20 and 6.21, the explicit expression based on Eqs. (6.45) and (6.46), and the approximate expression based on Eq. (6.48). All methods yield identical error standard deviations results to three significant figures. The empirical and explicit expression yield

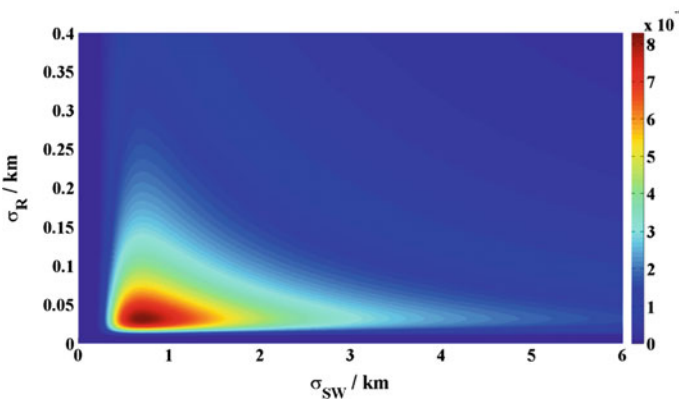


Fig. 6.20 Two-dimensional P_c distribution with color determined by magnitude of P_c

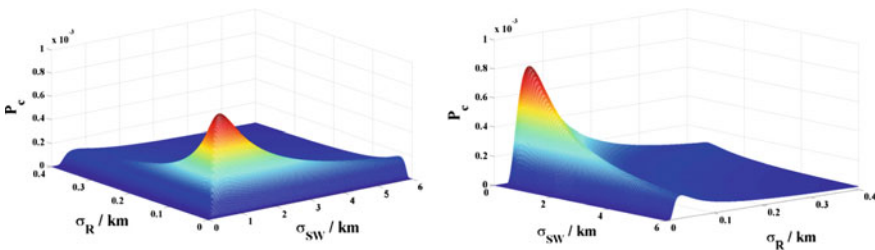


Fig. 6.21 Three-dimensional P_c mesh surface with respect to combined error standard deviations σ_R and σ_{SW} observed from different direction

Table 6.15 P_{cmax} and corresponding combined error standard deviations in arbitrary error ellipse shape case computed using empirical mesh surface, explicit expression, and approximate expression

Methods	$P_{cmax}/10^{-4}$	Relative error (%)	σ_R/km	σ_{SW}/km
Empirical	8.304292	—	0.031750	0.697500
Explicit expression	8.303965	0.0039	0.031748	0.697688
Approximate expression	8.313674	0.1	0.031730	0.697294

identical P_{cmax} results to three significant figures, while approximate expression’s result agrees to two significant figures.

3. Situation no. 5—variable orientation, fixed shape, fixed size

In this case we calculate P_{cmax} by Eq. (6.56). The resulted $P_{cmax} = 2.358194 \times 10^{-4}$. The corresponding angle which describes the orientation of error ellipse is $\theta = 0$ since $\sigma_x > \sigma_y$. To demonstrate P_{cmax} and the corresponding angle, we draw the curve

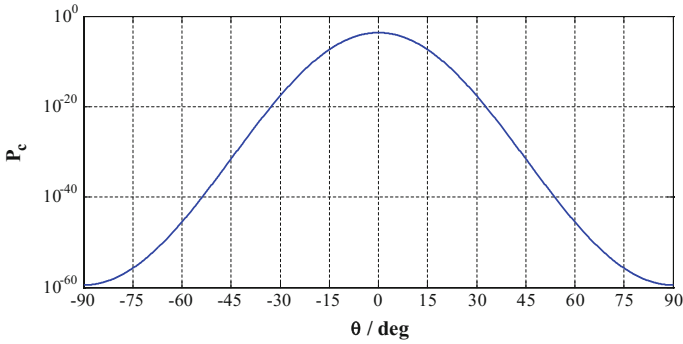


Fig. 6.22 The curve of P_c with respect to the angle θ (Situation no. 5)

Table 6.16 The P_{cmax} and corresponding combined error standard deviations (Situation no. 6)

P_{cmax}	5.154185×10^{-4}
σ_x/km	0.493744
σ_y/km	0.072279

of P_c with respect to the angle θ using Eq. (6.54) as illustrated in Fig. 6.22. The obtained empirical P_{cmax} is 2.358194×10^{-4} . The relative error of P_{cmax} is less than 3.8×10^{-7} .

4. Situation no. 6—variable orientation, fixed shape, variable size

In this case we calculate P_{cmax} and the corresponding combined error standard deviations by making use of Eq. (6.64). The resulted P_{cmax} and corresponding combined error standard deviations are shown in Table 6.16. The aspect ratio $\alpha = \sigma_x/\sigma_y = 6.831075$ is constant in this case.

To demonstrate P_{cmax} and the corresponding combined error standard deviations, we draw the curve of P_c with respect to combined error standard deviations σ_x using Eq. (6.58) as illustrated in Fig. 6.23. The empirical P_{cmax} and corresponding combined error standard deviations obtained from Fig. 6.23 are $P_{cmax} = 5.154185 \times 10^{-4}$ and $\sigma_x = 0.493744$ km, $\sigma_y = 0.072279$ km. The relative error of P_{cmax} is less than 1.7×10^{-7} .

5. Situation no. 8—variable orientation, variable shape, variable size

In this case we calculate P_{cmax} and the corresponding combined error standard deviations by making use of Eqs. (6.70) and (6.73). The resulted P_{cmax} and corresponding combined error standard deviations are shown in Table 6.17. Note that in this case $\sigma_y \rightarrow 0$.

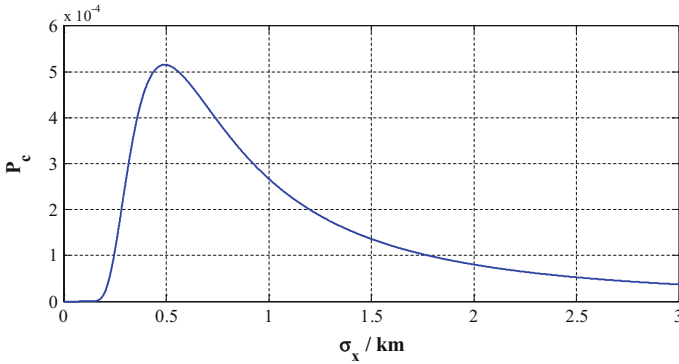


Fig. 6.23 The curve of P_c with respect to combined error standard deviations σ_x (Situation no. 6)

Table 6.17 The P_{cmax} and corresponding combined error standard deviations (Situation no. 8)

P_{cmax}	6.933103×10^{-3}
σ_x / km	0.697992
σ_y / km	0

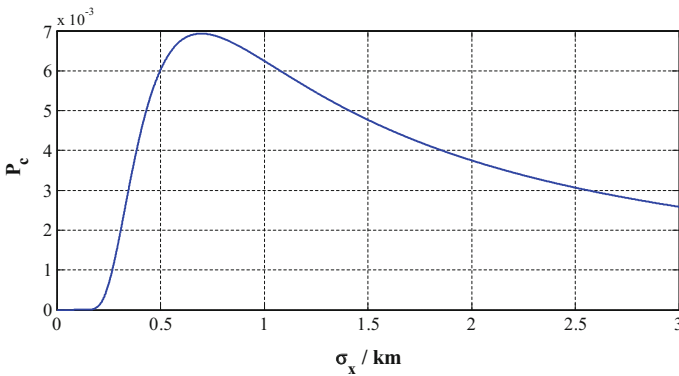


Fig. 6.24 The curve of P_c with respect to combined error standard deviations σ_x (Situation no. 8)

To demonstrate P_{cmax} and the corresponding combined error standard deviations, we draw the curve of P_c with respect to combined error standard deviations σ_x using Eq. (6.67) as illustrated in Fig. 6.24. We can obtain empirical P_{cmax} and corresponding combined error standard deviations from Fig. 6.24. The empirical P_{cmax} is 6.933103×10^{-3} and $\sigma_x = 0.697992$ km. The relative error of P_{cmax} is less than 1.3×10^{-7} .

6.2.5 Discussion

Table 6.18 summarizes the resulted P_{cmax} and corresponding error covariance (scaling or orientating factor) for five situations. Figure 6.25 shows the bar of P_c and P_{cmax} of five situations. Table 6.18 and Fig. 6.25 indicate that:

- (1) P_c in situation No.1 is the actual P_c for the given set of close approach data and error covariance data, so it is the smallest one in all of the situations.
- (2) P_{cmax} in situation No. 8 (P_{cmax8}) is the largest one in all of the situations, this result is predictable because situation No. 8 is the “worst” situation in which the orientation, shape, and size of covariance ellipse are variables. The P_{cmax8} is the “absolute” maximum probability.
- (3) Both the situations No. 2 and No. 4 are fixed orientation cases, so they are comparable. P_{cmax4} is larger than P_{cmax2} because in situation No. 2 only the size is variable while in situation No. 4 not only the size but also the shape is a variable. Likewise, in situations No. 5, 6 and 8 the orientations of covariance

Table 6.18 Summary of the resulted P_{cmax} and corresponding error covariance of five situations

Situation	Situation description	Maximum probability	Corresponding error covariance (Scaling or orientating factor)
1	Fixed orientation	1.807912×10^{-4}	$\sigma_x = 0.2941297, \sigma_y = 0.0430576$
	Fixed shape		
	Fixed size		
2	Fixed orientation	4.710037×10^{-4}	$k = 1.756027$
	Fixed shape		
	Variable size		
4	Fixed orientation	8.303965×10^{-4}	$\sigma_x = 0.697688, \sigma_y = 0.031748$
	Variable shape		
	Variable size		
5	Variable orientation	2.358194×10^{-4}	$\theta = 0(\sigma_x > \sigma_y)$
	Fixed shape		
	Fixed size		
6	Variable orientation	5.154185×10^{-4}	$\sigma_x = 0.493744, \sigma_y = 0.072279$
	Fixed shape		
	Variable size		
8	Variable orientation	6.933103×10^{-3}	$\sigma_x = 0.697992, \sigma_y = 0$
	Variable shape		
	Variable size		

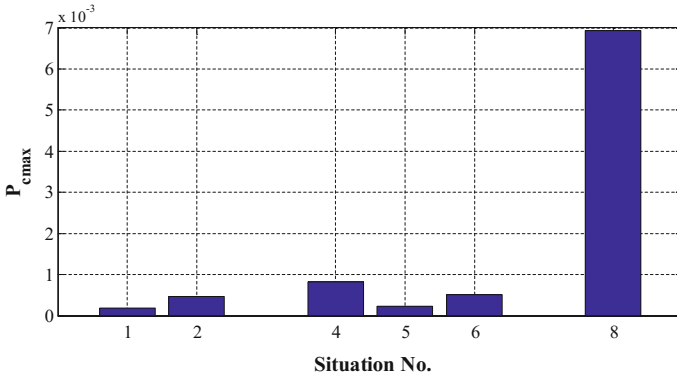


Fig. 6.25 The P_c and P_{cmax} of five situations

ellipse are variable, and in situation No. 5 the shape and size are both fixed, in situation No. 6 the shape is fixed but the size is variable, in situation No. 8 the shape and size are both variables, hence P_{cmax8} is larger than P_{cmax6} and P_{cmax6} is larger than P_{cmax5} .

- (4) In this conjunction case the angle θ is small (2.6055° , see Fig. 5.12) and close to $\theta = 0$ which maximizes P_c when the orientation is variable. Therefore the variability of the orientation has little influence on P_{cmax} , while the variability of the shape and the size influence much on P_{cmax} . Therefore, P_{cmax4} is larger than P_{cmax5} and P_{cmax6} .
- (5) P_c is close to P_{cmax5} , and P_{cmax2} is close to P_{cmax6} , as the angle θ is close to zero. A mild rotation of covariance ellipse can transform situation from No. 1 to No. 5, and from situation No. 2 to No. 6. Note that P_{cmax5} is larger than P_c and P_{cmax6} is larger than P_{cmax2} .

Generally speaking, the situation No. 2 where only the size is variable is closest to the reality (actual P_c). The reason why P_{cmax5} is less than P_{cmax2} in this example is that the angle θ approached zero, which is not common in conjunction assessment. The “worst” situation No. 8 where all the orientation, shape, and size of covariance ellipse are variables is farthest from the reality. In practical operations, one should fix as many factors (orientation, shape, and size) as possible to approximate the actual P_c . The less P_{cmax} is, the more meaningful it is.

In addition, one should be aware that P_{cmax} can differ drastically from the actual P_c for a given set of close approach data. The examples presented in this section indicate that P_c varies slightly from P_{cmax} because of the relatively small covariances. The error standard deviations shown in Table 6 are dozens of meters in R and W direction and hundreds of meters in S direction. If the covariances are enlarged to 10 times, the error in S direction is about several kilometers, and P_c will be 3.828153×10^{-5} . In this case P_{cmax} is much larger than the actual value of P_c , which is more common in the preliminary conjunction analysis.

Generally, P_{cmax} might differ by orders of magnitude from the actual value computed directly from the data. P_{cmax} is a one-sided test. If P_{cmax} is less than an acceptable level it is useful (no action required). The definition of the action threshold of P_{cmax} is up to the operator since the conjunction assessment is not totally automatic at present. To avoid missing alarm, the threshold of P_{cmax} should be less than the least P_c threshold (yellow threshold for instance). If P_{cmax} is greater than the action threshold it makes no sense because the actual computed P_c might be below the threshold by orders of magnitude. More observations and refined collision probability analysis are necessary in this case.

Compared with previous works, this section provides a clear and comprehensive discussion to address all the situations when assessing the maximum collision probability. The discussion is based on Chan's analytical equation for calculating the collision probability. First, all situations when calculating P_{cmax} are categorized according to the changeability of the orientation, shape, and size of covariance ellipse. Eight situations have been obtained. One of them does not yield P_{cmax} because there is no variable factor, and two of them are not practical in which the shapes of covariance ellipse are variable but the sizes are fixed. The other five situations are discussed with uniform definitions and symbols. The consequences are compared to and validated by the ones yielded from previous works. Finally, a practical conjunction event as a test case is presented to demonstrate and validate the results.

The question regarding which method is the best for conjunction assessment needs to be answered. Actually, there is no best solution but the most suitable one. Which method should be used for conjunction assessment depends on what are the known conjunction conditions (the conjunction geometry, the shape, and size of the covariances), and their confidence level. It is up to the operator to decide which is best for their purposes based on the mission requirement, the known condition, and their accuracies.

6.3 Missing- and False-Alarm Analysis of Conjunction Assessment

The collision probability (P_c) between two objects in the conjunction is one of the most important criteria for conjunction risk assessment. The calculation of P_c , using information of position and velocity vectors and associated error covariance at the TCA, is one of the key techniques in conjunction risk assessment, and has been well studied. The P_c is closely relative to the position and velocity vectors, associated error covariance, and combined radius of two space. These informations are obtained via observations, estimations, or computations. The inherent uncertainty of these informations will affect the confidence level of conjunction risk assessment. The P_c is calculated for risk assessment and collision avoidance decision-making. How to make decision correctly using P_c in practical operations is an important

issue. Consequently, it is not sufficient to just obtain the risk assessment parameters such as the closest approach distance and collision probability. The comprehensive assessment of collision risk based on the calculate P_c is requisite to improve the effectiveness and confidence level.

In this section, the safety-region and danger-region of conjunction assessment are defined by using the explicit expression of P_c , the methods for calculating probabilities of missing alarm and false alarm are presented. In this way, it is useful to directly know the reliability of P_c , balance the cost of taking action to maneuver the spacecraft or not, and make the final decision.

6.3.1 Essential of Conjunction Assessment and Misjudgment

The basic procedure of conjunction assessment is as follow. Collision risk parameters (such as closest approach distance, collision probability, maximum collision probability) are obtained from collision risk assessment by using the predicted orbital states and covariance informations. The risk parameters are judged according to certain criteria whether they are in the danger-region. For example, if miss distance is less than the distance threshold, or collision probability is greater than the probability threshold. The collision alarm will be given and appropriate measures should be taken if the risk parameters are in danger-region. Otherwise, two space objects are supposed to be safe.

The criteria for decision-making in probability-based conjunction assessment are, if the calculated P_c is greater than the pre-defined probability threshold P_T ($P_c > P_T$), this event will be determined as “danger”; if the P_c is less than the probability threshold P_T ($P_c < P_T$), this event will be determined as “safety”.

Conjunction assessment is essentially a discriminant analysis problem. As a discriminant problem, there are always two types of misjudgment: missing alarm and false alarm. Missing alarm means two space objects will actually collide but we suppose they are safety. False alarm means two space objects are safety but we suppose they are dangerous to collide. Accordingly, there are two kinds of probability: missing alarm probability P_m and false alarm probability P_{fa} . The missing- and false-alarm will affect the confidence level.

The missing- and false-alarm analysis is an important issue in discriminant analysis. Both the probability of missing alarm (P_m) and the probability of false alarm (P_{fa}) are needed to be controlled in a conjunction risk assessment, but they conflict to each other and need to be balanced. In general, the smaller P_m the greater P_{fa} is, and the greater P_m the smaller P_{fa} is.

The conjunction assessment is affected by all kinds of factors, and the decision is making with all kinds of errors. The missing- and false-alarm are inevitable and will influence the confidence level of conjunction assessment. This section provides the concepts of probabilities of missing alarm and false alarm in a conjunction event.

Based on the explicit expression of collision probability, the safety-region and danger-region of conjunction assessment are defined, the methods for calculating probabilities of missing alarm and false alarm are put forward, and the basic property of probabilities of missing alarm and false alarm are provided.

Analysis of missing alarm and false alarm is still an exploratory research in which there are many issues that need to be resolved. In this section, a preliminary analysis of this issue is given out, which provides ideas for further research.

6.3.2 Safety- and Danger-Region of Conjunction Assessment

As stated before, the false alarm in conjunction assessment is that the risk assessment parameter fall in the danger region then leads to alarm and maneuver while the collision will not actually happen; the missing alarm is that the risk assessment parameters fall in the safety-region then do not lead to alarm and maneuver while the collision will happen. The predicted state (position and velocity) at TCA used to calculate P_c is not the “true” state. Actually, the true state falls in a multidimensional ellipsoid centered at the predicted state with high probability and described by the covariance matrix. For a conjunction event, the true state is determinate but unknown, the predicted states from different kinds of observations and determination algorithms distributed statistically around the true state. Based on this understanding, the safety- and danger-region, missing- and false-alarm of conjunction assessment can be determined.

Assuming the actual relative position vector on the conjunction plane is $(x_{\text{true}}, y_{\text{true}})$, and the predicted relative position vector is $(x_{\text{pred}}, y_{\text{pred}})$, the predicted vector is distributed around the true state represented by Gaussian distribution

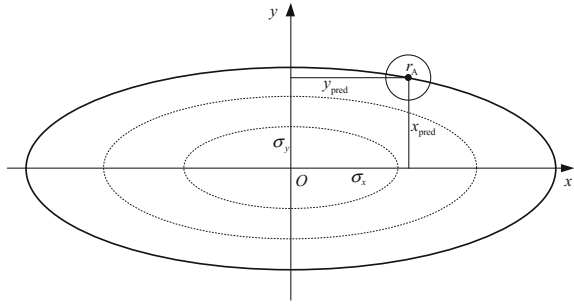
$$x_{\text{pred}} \sim N(x_{\text{true}}, \sigma_x^2), \quad y_{\text{pred}} \sim N(y_{\text{true}}, \sigma_y^2) \quad (6.75)$$

Collision probability is calculated by the prediction relative position vector $(x_{\text{pred}}, y_{\text{pred}})$, which can be explicit expressed as

$$P_c = \exp \left[-\frac{1}{2} \left(\frac{x_{\text{pred}}^2}{\sigma_x^2} + \frac{y_{\text{pred}}^2}{\sigma_y^2} \right) \right] \left[1 - \exp \left(-\frac{r_A^2}{2\sigma_x\sigma_y} \right) \right] \quad (6.76)$$

Since the prediction state $(x_{\text{pred}}, y_{\text{pred}})$ is random, the calculated collision probability P_c is also random. Equation (6.76) could define the ellipses of constant collision probability as shown in Fig. 6.26, the ellipses can be formulated as

Fig. 6.26 Ellipses of constant collision probability



$$\frac{x_{\text{pred}}^2}{\sigma_x^2} + \frac{y_{\text{pred}}^2}{\sigma_y^2} = C^2, \quad C > 0 \tag{6.77}$$

For those predicted states on the ellipse, the values of calculated collision probability are constant.

Let the alarm threshold of collision probability is P_T , the safety-region S and the danger-region D of conjunction assessment can be expressed as follows

$$\begin{cases} S : \{ (x_{\text{pred}}, y_{\text{pred}}) | P_c(x_{\text{pred}}, y_{\text{pred}}) < P_T \} \\ D : \{ (x_{\text{pred}}, y_{\text{pred}}) | P_c(x_{\text{pred}}, y_{\text{pred}}) \geq P_T \} \end{cases} \tag{6.78}$$

From Eq. (6.76), the safety-region S can be written as

$$\begin{aligned} P_c &= \exp \left[-\frac{1}{2} \left(\frac{x_{\text{pred}}^2}{\sigma_x^2} + \frac{y_{\text{pred}}^2}{\sigma_y^2} \right) \right] \left[1 - \exp \left(-\frac{r_A^2}{2\sigma_x\sigma_y} \right) \right] < P_T \\ \Leftrightarrow \frac{x_{\text{pred}}^2}{\sigma_x^2} + \frac{y_{\text{pred}}^2}{\sigma_y^2} &> 2 \left[\ln \left(1 - e^{-\frac{r_A^2}{2\sigma_x\sigma_y}} \right) - \ln P_T \right] \end{aligned} \tag{6.79}$$

Let

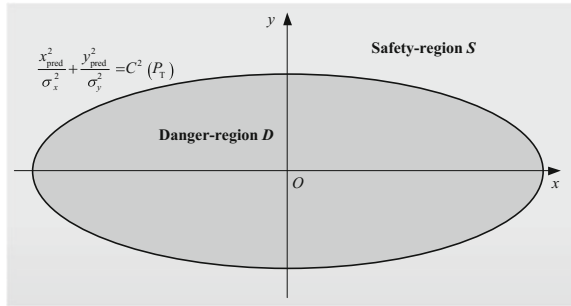
$$C^2(P_T) = 2 \left[\ln \left(1 - e^{-\frac{r_A^2}{2\sigma_x\sigma_y}} \right) - \ln P_T \right] \tag{6.80}$$

Thus the S region is the region outside the following ellipse (the gray infinite region in Fig. 6.27) when the alarm threshold of collision probability is P_T

$$\frac{x_{\text{pred}}^2}{\sigma_x^2} + \frac{y_{\text{pred}}^2}{\sigma_y^2} = C^2(P_T) \tag{6.81}$$

The D region is inside above ellipse (the dark ellipse region in Fig. 6.27).

Fig. 6.27 Safety- and danger-region of conjunction assessment



Then the safety-region S and danger-region D can be expressed as

$$\begin{cases} S : \left\{ (x_{\text{pred}}, y_{\text{pred}}) \mid \frac{x_{\text{pred}}^2}{\sigma_x^2} + \frac{y_{\text{pred}}^2}{\sigma_y^2} > C^2(P_T) \right\} \\ D : \left\{ (x_{\text{pred}}, y_{\text{pred}}) \mid \frac{x_{\text{pred}}^2}{\sigma_x^2} + \frac{y_{\text{pred}}^2}{\sigma_y^2} \leq C^2(P_T) \right\} \end{cases} \quad (6.82)$$

Equation (6.81) indicates that the size of boundary ellipse between S and D region is determined by the probability threshold P_T . When P_T is growing, the boundary ellipse is becoming smaller and the S region is becoming larger. When P_T is decreasing, the boundary ellipse is becoming larger and the S region is becoming smaller.

Let $r_A = 20$, $\sigma_x = 1000$, $\sigma_y = 100$ (distance unit), Fig. 6.28 illustrates the curve of the size of boundary of ellipse C versus the collision probability threshold P_T . When P_T is greater than a certain threshold value, the size of boundary ellipse C is equal to 0, which indicates that the danger-region is degraded to the coordinate origin, and the entire two-dimensional conjunction plane is the safety-region. That means the collision probability threshold is greater than the maximum collision

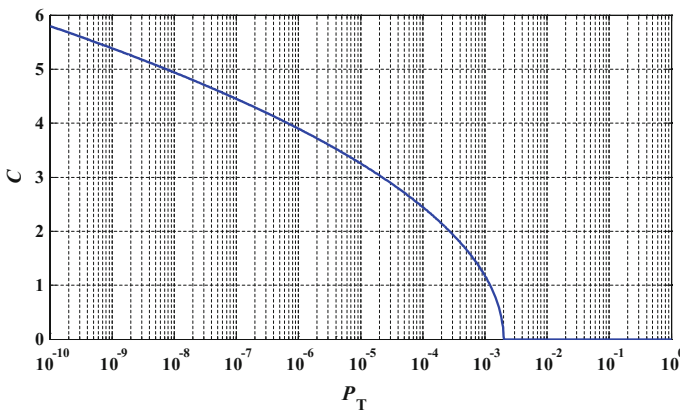


Fig. 6.28 Curve of size of boundary ellipse versus probability threshold P_T

probability. The collision probability in any case will be less than the collision probability threshold. The conjunction risk assessment is always determined as safety.

6.3.3 Calculation of Probability of Missing Alarm

The probability of missing alarm P_m is defined as the probability that two space objects will actually collide but supposed to be safety. P_m equals to the probability of predicted relative position vector (x_{pred}, y_{pred}) falls in the safety-region S when the actual relative position vector (x_{true}, y_{true}) falls in the combined hardball of two space objects $(x_{true}^2 + y_{true}^2 \leq r_A^2)$. Then, P_m can be represented as

$$P_m = P((x_{pred}, y_{pred}) \in S | x_{true}^2 + y_{true}^2 \leq r_A^2) \tag{6.83}$$

As shown in Fig. 6.29, the actual relative position vector (x_{true}, y_{true}) falls in the collision circular area (its radius is r_A), the predicted relative position vector (x_{pred}, y_{pred}) is distributed around the actual relative position vector satisfying Gaussian distribution. The dashed ellipse in Fig. 6.29 is the 3σ covariance ellipse. Note that the missing alarm region is spreading the entire two-dimensional safety-region. The missing alarm region within the 3σ covariance ellipse is shown to characterize.

Equation (6.75) reveals that $(x_{pred}^2/\sigma_x^2) + (y_{pred}^2/\sigma_y^2)$ satisfies the non-central χ^2 distribution with 2 degrees of freedom [17]

$$\frac{x_{pred}^2}{\sigma_x^2} + \frac{y_{pred}^2}{\sigma_y^2} \sim \chi^2(2, \delta), \quad \delta = \sqrt{\frac{x_{true}^2}{\sigma_x^2} + \frac{y_{true}^2}{\sigma_y^2}} \tag{6.84}$$

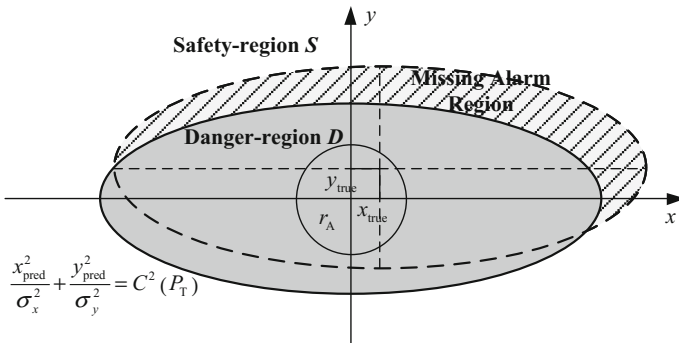


Fig. 6.29 Missing alarm region of collision assessment

where δ is the non-central parameter. Let the probability distribution function of the non-central χ^2 distribution is $F(x|2, \delta)$, the probability of missing alarm can be expressed as

$$P_m = 1 - F(C^2|2, \delta), \quad \delta = \sqrt{\frac{x_{\text{true}}^2}{\sigma_x^2} + \frac{y_{\text{true}}^2}{\sigma_y^2}}, \quad x_{\text{true}}^2 + y_{\text{true}}^2 \leq r_A^2 \quad (6.85)$$

Equation (6.85) indicates that probability of missing alarm P_m is determined by not only the size of boundary ellipse C (hence the threshold P_T), but also the position of actual relative position vector $(x_{\text{true}}, y_{\text{true}})$. Let $r_A = 20$, $\sigma_x = 1000$, $\sigma_y = 100$ (distance unit), $P_T = 10^{-4}$. The size of the boundary ellipse $C = 2.447$ can be yielded by Eq. (6.80). Figure 6.31 shows the curve surface of the missing alarm probability P_m changing with the actual relative position $(x_{\text{true}}, y_{\text{true}})$ in the collision plane.

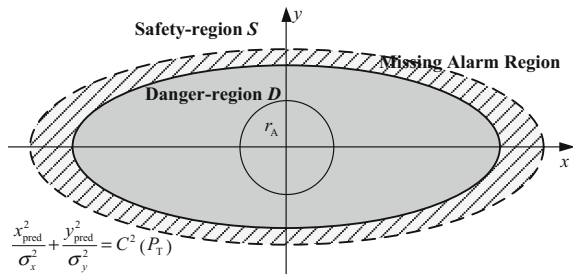
As shown in Fig. 6.31, the missing alarm probability P_m maximized when the actual relative position $(x_{\text{true}}, y_{\text{true}})$ is at the end of minor axis direction (y direction in this case). In this case, the maximum probability of missing alarm $P_{m,\text{max}} = 0.0654$ when $(x_{\text{true}}, y_{\text{true}})$ is at the two separate points of $(0, 20)$ and $(0, -20)$. The expression of the maximum probability of missing alarm is

$$P_{m,\text{max}} = 1 - F(C^2|2, \delta), \quad \delta = \frac{r_A}{\sigma_y} \quad (6.86)$$

The size of boundary ellipse $C = 2.447$ when the alarm threshold of collision probability is $P_T = 10^{-4}$, which means the boundary ellipse is actually the 2.447σ covariance ellipse. The collision circular area is small compared to the boundary ellipse. It is reasonable to approximately suppose that the actual relative position is at the coordinate origin $(x_{\text{true}} = y_{\text{true}} = 0)$. Therefore the region of missing alarm becomes an elliptical ring as shown in Fig. 6.30. The missing alarm probability in this case can be simplified to

$$P_m = \frac{P_T}{1 - \exp\left(-\frac{r_A^2}{2\sigma_x\sigma_y}\right)} \quad (6.87)$$

Fig. 6.30 Missing alarm region when the true value is in the origin



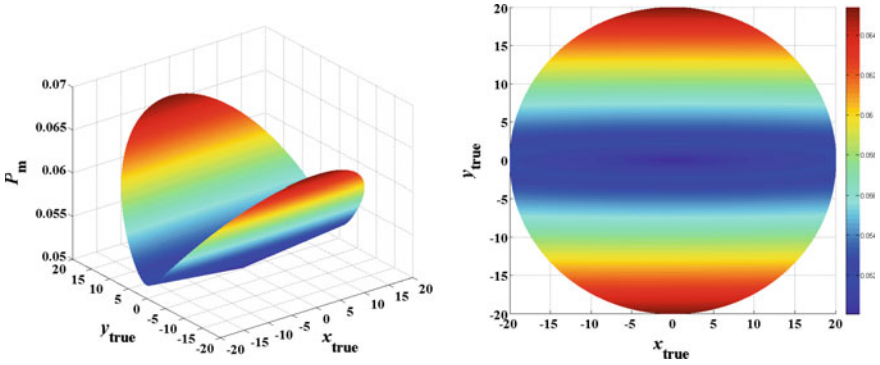


Fig. 6.31 Curve surface of P_m versus (x_{true}, y_{true})

Let $r_A = 20$, $\sigma_x = 1000$, $\sigma_y = 100$ (distance unit), and $P_T = 10^{-4}$, then the missing alarm probability is $P_m = 0.05005$, which is identical to the P_m at origin in Fig. 6.31. Figure 6.31 reveals that the P_m at origin is the minimum probability. The distinction between the maximum and minimum P_m is not inapparent because the collision circular area is small. Equation (6.87) can be used as approximate formula of missing alarm probability in preliminary analysis.

When the following inequality satisfied

$$\frac{P_T}{1 - \exp\left(-\frac{r_A^2}{2\sigma_x\sigma_y}\right)} > 1 \tag{6.88}$$

the missing alarm probability is always equal to one ($P_m = 1$).

$$P_T > 1 - \exp\left(-\frac{r_A^2}{2\sigma_x\sigma_y}\right), \quad \sigma_x\sigma_y > \frac{r_A^2}{-2 \ln(1 - P_T)} \tag{6.89}$$

Equation (6.89) means that it is always missing alarm when the alarm threshold of collision probability P_T or the orbit propagation errors are overlarge. That is, the maximum collision probability is smaller than the threshold of collision probability when Eq. (6.89) is satisfied.

6.3.4 Calculation of Probability of False Alarm

The probability of false alarm P_{fa} is defined as the probability that two space objects will not collide but supposed to be danger. P_{fa} equals to the probability of predicted relative position vector (x_{pred}, y_{pred}) falls in the danger-region D when the actual

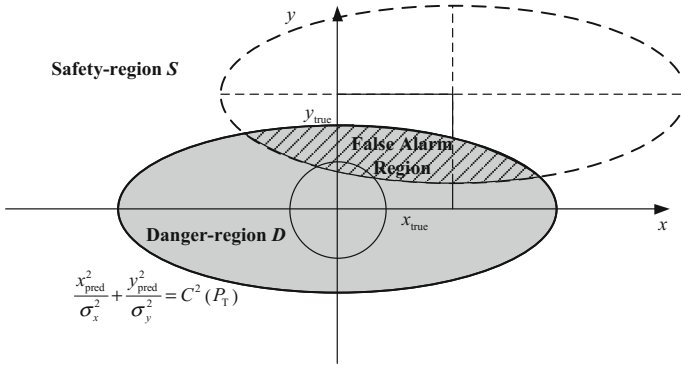


Fig. 6.32 False alarm region of collision assessment

relative position vector $(x_{\text{pred}}, y_{\text{pred}})$ falls outside the combined hardball of two space objects $(x_{\text{true}}^2 + y_{\text{true}}^2 > r_A^2)$. Then, P_{fa} can be represented as

$$P_{\text{fa}} = P((x_{\text{pred}}, y_{\text{pred}}) \in D | x_{\text{true}}^2 + y_{\text{true}}^2 > r_A^2) \tag{6.90}$$

As shown in Fig. 6.32, the actual relative position vector $(x_{\text{true}}, y_{\text{true}})$ falls outside the collision circular area (its radius is r_A), the predicted relative position vector $(x_{\text{pred}}, y_{\text{pred}})$ is distributed around the actual relative position vector satisfying Gaussian distribution. The dashed ellipse in Fig. 6.32 is the 3σ covariance ellipse. Note that the false alarm region is the whole danger-region. The missing alarm region within the 3σ covariance ellipse is shown to characterize.

Similar to the missing alarm probability, the false alarm probability can also be described by the non-central χ^2 distribution with 2 degrees of freedom. The expression of false alarm probability is

$$P_{\text{fa}} = F(C^2 | 2, \delta), \quad \delta = \sqrt{\frac{x_{\text{true}}^2}{\sigma_x^2} + \frac{y_{\text{true}}^2}{\sigma_y^2}}, \quad x_{\text{true}}^2 + y_{\text{true}}^2 > r_A^2 \tag{6.91}$$

where δ is the non-central parameter. Equation (6.91) indicates that probability of false alarm P_{fa} is determined by not only the size of boundary ellipse C (hence the threshold P_T) but also the position of actual relative position vector $(x_{\text{true}}, y_{\text{true}})$. Let $r_A = 20$, $\sigma_x = 1000$, $\sigma_y = 100$ (distance unit), and $P_T = 10^{-4}$. The size of the boundary ellipse $C = 2.447$ can be yielded by Eq. (6.80). Figure 6.33 shows the curve surface of the false alarm probability P_{fa} changing with the actual relative position $(x_{\text{true}}, y_{\text{true}})$ in the collision plane but outside the collision circular area. The area outside the collision circular is infinite, Fig. 6.33 illustrates it up to $\sqrt{x_{\text{true}}^2 + y_{\text{true}}^2} = 2000$.

As shown in Fig. 6.33, the false alarm probability P_{fa} is maximum when the actual relative position $(x_{\text{true}}, y_{\text{true}})$ is at the end of major axis direction (x direction

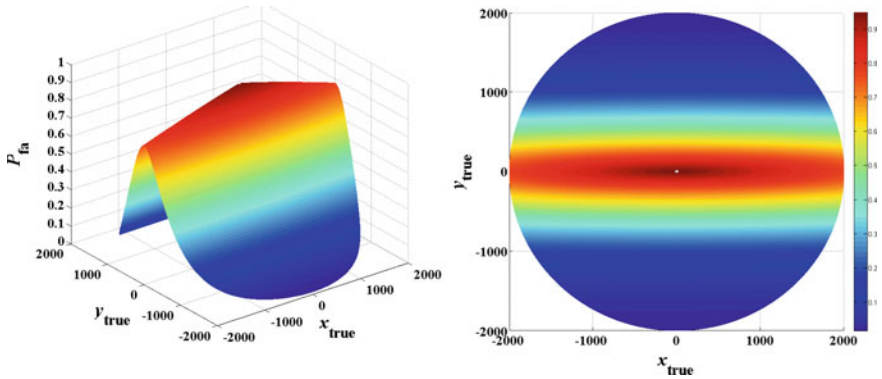


Fig. 6.33 Curve surface of P_{fa} versus (x_{true}, y_{true})

in this case). In this case, the maximum probability of false alarm $P_{fa,max} = 0.9484$ when (x_{true}, y_{true}) is at the two separate points of $(-20, 0)$ and $(20, 0)$. The expression of the maximum probability of false alarm is

$$P_{fa,max} = F(C^2 | 2, \delta), \quad \delta = \frac{r_A}{\sigma_x} \tag{6.92}$$

Figure 6.33 also shows that the false alarm probability P_{fa} decreases slowly as the actual relative position (x_{true}, y_{true}) changes along the major axis of the error ellipse (x direction) in the collision plane, which is equal to 0.77 when (x_{true}, y_{true}) is already at $(\pm 2000, 0)$. But it decreases quickly as (x_{true}, y_{true}) is changing along the minor axis (y direction), which is equal to 0.0148 when (x_{true}, y_{true}) is at $(0, \pm 2000)$.

6.4 Comprehensive Assessment of Collision Risk

The conjunction risk assessment is essentially a transform from the predicted conjunction parameters to the risk information could be used in the operations. This process is a quantitative analysis of collision risk of the conjunction event, and determining how high the confidence level of the quantitative risk is. Presently, the miss distance and collision probability (P_c) are widely used to evaluate the risk of collision. Individual metrics, such as the miss distance and collision probability (P_c), do not singly capture the overall risk associated with a conjunction.

In miss distance analysis, regional approaching area defined by RSW coordinate around one of two space objects (primary object) is used to determine collision risk between two space objects. When the other object (secondary object) falls into the threshold of RSW region, the approaching event is considered to be dangerous, otherwise safe. The method of determination is “1” and “0” in practice, in which the

minimum approaching distance is the sole criterion of collision risk. That means it is very simple and can be easily used.

Compared to the miss distance, collision probability between two space objects is easier to understand, but more complicated to calculate. It is determined by the relative position and velocity vectors, relative position covariance and the combined radius of two objects which are considered as sphere. These parameters are inevitably imprecise as they are all acquired by measurement, estimation and calculation. The accuracy of these parameters will affect the reliability of P_c . For example, the P_c could be indicating a risk, but a high P_c may be due to unrealistic covariance sizes due to insufficient tracking.

Both the miss distance and P_c have limitations to give out authentic risk of collision, as they are only one aspect used as conjunction risk parameters. A comprehensive assessment of collision risk considering all kinds of factors is necessary. For a conjunction event, the factors should be considered in the comprehensive assessment of collision risk consist of risk assessment parameters and quality assessment parameters. The risk assessment parameters, such as miss distance and P_c , are used to describe the risk level. The quality assessment parameters are used to describe the accuracy and reliability of data, such as days to the time of closest approach (TCA), days since the last observation, orbit prediction errors, and so on.

It is difficult to carry out an overall assessment only by risk assessment parameters or quality assessment parameters. For example, when miss distance falls into the threshold of distance but P_c is beyond the threshold of collision probability, or the same P_c comes from different orbit propagation errors, how to assess these risk of collision? Therefore, risk assessment parameters and quality assessment parameters should be considered simultaneously in a conjunction event.

These parameters contain sufficient informations. But they are complex for decision-maker. While conjunction assessment is admittedly a complex problem, being able to convey the risk level in a simple manner is something that has been expressed as very desirable by mission stakeholders. Aim at this requirement, the comprehensive assessment of conjunction risk provides a single metric that can serve as a risk level index which quantifies the overall risk associated with a conjunction.

6.4.1 Definition of the *F-Value*

Frigm of NASA Goddard Space Flight Center has put forward the concept of a single risk metric *F-value* to estimate the conjunction risk between space objects [18]. The concept of the *F-value* is an extension of Fuzzy Set Theory and Fuzzy Logic, which conveys a succinct risk level by combining miss distance, collision probability and orbit determination solution into one single value, and provides a simple and consistent metric of the comparison of conjunction risks.

The F -value for a particular conjunction combines any desired number of conjunction risk and quality assessment parameters into one single value. A risk assessment parameter is any quantifiable metric that is used in dispositioning the risk associated with a given conjunction, such as miss distance and probability of collision. A quality assessment parameter is any quantifiable metric that is used in assessing the quality of the orbit determination solution off of which the close approach predictions are based.

The basic calculating model of the F -value is defined as Eq. (6.93), which means the overall F -value for a conjunction event is the product of a weighting average of the f -value of all risk assessment parameters and a weighting average of the f -value of all the quality assessment parameters [18].

$$F = \left(\frac{1}{m} \sum_{i=1}^m a_i f_i \right) \cdot \left(\frac{1}{n} \sum_{j=1}^n b_j f_j \right) \quad (6.93)$$

where m is the number of risk assessment parameters, n is the number of quality assessment parameters, f is the specific f -value of each parameters mapped by the membership functions, a and b are the weighting coefficients of each parameter.

A specific f -value is a reduction of the actual value of the parameter to a common scale. In other words, an f -value maps the range of all possible values for a particular parameter to a common range of values using a membership function. Once the values have been re-scaled, the aggregation becomes possible. An F -value only considering the risk assessment parameters, and not the quality assessment parameters, is also referred to as the “unscaled” F -value. Conversely, an F -value including both sets of parameters is the “scaled” F -value or, simply, the F -value.

There are several ways to tune the F -value to meet any specific conjunction assessment requirements or desired preferences. Specifically, there are three tuning mechanisms: the risk or quality assessment parameters used, relationship between an assessment parameter value and associated f -value (i.e. membership function); weighting of each assessment parameter.

The Eq. (6.93) used to calculate F -value is relatively simple, easy to implement and more flexible. The membership functions of f , weighting coefficients a and b can be adjusted by historical conjunction events.

6.4.2 Risk- and Quality-Assessment Parameters

The risk assessment parameters include the miss distance and collision probability. The miss distance can be the range between two objects or the radial (R), along-track (S), and cross-track (W) components of the relative position. Frigm used five risk assessment parameters including the collision probability, miss distance (Root Sum Square), the radial miss distance, along-track miss distance and cross-track miss distance [18]. Actually, the miss distance and its components are

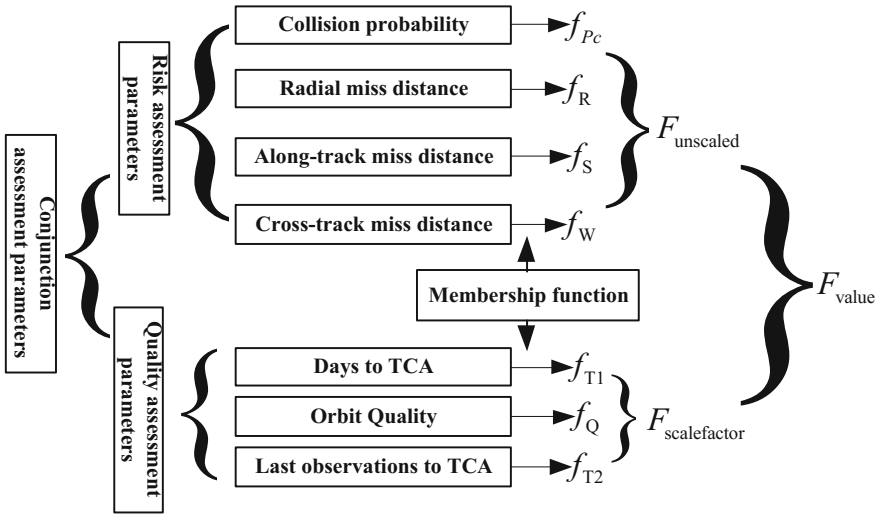


Fig. 6.34 Comprehensive assessment of collision risk based on the F -value

not independent. The R, S, W components are the projections of miss distance to radial, along-track and cross-track directions. When the miss distance and its RSW components are considered simultaneously in F -value calculation, it may increase the weight of the distance factor. Thus, we only consider the collision probability and RSW components as four risk assessment parameters.

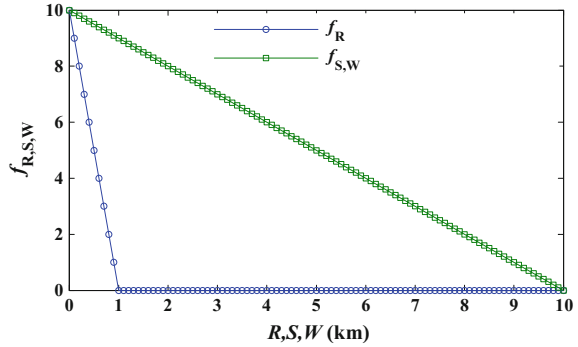
The quality assessment parameters include the effect of time and accuracy of orbit determination. Frigm also used three quality assessment parameters, including days to TCA, determinant of the combined positional covariance and days since last observation of objects [18]. As the collision probability calculation has taken into account the effect of the covariance matrix, and the determinant of the combined covariance cannot fully reflect the orbit determining and propagating accuracy, we use days to TCA, quality of orbit determination, and last observations to TCA as quality assessment parameters.

The quality of orbit determination ranges from 0 to 10, 0 means the worst and 10 means the best. This parameter is rated by experts considering the accuracy of observations and the orbit determination model. The comprehensive assessment of collision risk based on the F -value is shown in Fig. 6.34.

6.4.3 Membership Functions of Assessment Parameters

Both the dimensions and values of different assessment parameters are distinct evidently. For example, the collision probability is generally below 10^{-3} , the miss distance may be 10 m to several kilometers, the time parameters may be several

Fig. 6.35 Membership function curve of miss distance



days to dozens of days, the quality of orbit determination ranges from 0 to 10. They are not comparable. Therefore, the membership functions, whose output has uniform unit and values, are chosen to construct the relationships between parameters and f values. In Fuzzy Set Theory, the most common membership function is linear and curve. In the section, the f value of risk assessment parameter ranges from 0 to 10, the f value of quality assessment parameter ranges from 0 to 1.

1. Membership function of miss distance

The membership function of RSW miss distance is chosen to be linear. Assuming D denotes the absolute value of components of miss distance (radial, along-track and cross-track); the membership function can be expressed as

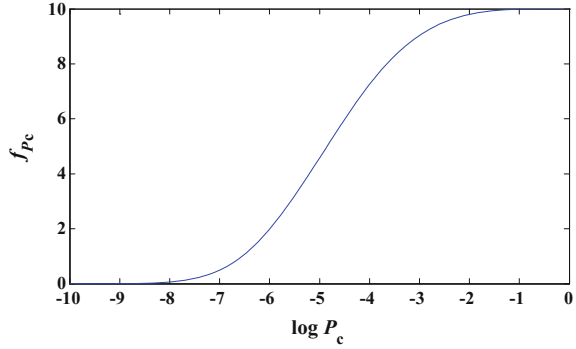
$$f_D = \begin{cases} \frac{10(L_D - D)}{L_D} & 0 \leq D \leq L_D \\ 0 & D > L_D \end{cases} \quad (6.94)$$

The parameter L_D of three components is different. The explicit expression of collision probability in Chap. 5 reveals that the position errors in the horizontal plane (S and W directions) are coupled, and related to the horizontal components of relative position at conjunction. The position errors in the radial direction (R) are decoupled with the position error in the horizontal plane, and related to the radial component of the relative position at conjunction. As the orbit prediction error in R direction is relative small, the values of L_R should be smaller. The S and W direction distance are coupled, so L_S and L_W should be equal and bigger. Therefore, $L_R = 1$ km, $L_S = L_W = 10$ km. The membership function curve of RSW miss distance is shown in Fig. 6.35.

2. Membership function of collision probability

The membership function of collision probability is chosen to be curvilinear, which can be expressed as

Fig. 6.36 Membership function curve of collision probability



$$f_{P_c} = \begin{cases} 10e^{-k(\log P_c)^4} & L_{P_c} \leq P_c \leq 1 \\ 0 & P_c < L_{P_c} \end{cases} \quad (6.95)$$

Let $k = 0.00125$, $L_{P_c} = 10^{-10}$, the membership function curve of collision probability is shown in Fig. 6.36. The curve increases slowly at first with P_c values around 10^{-10} , 10^{-9} , and 10^{-8} , then increases rapidly with P_c values around 10^{-6} , 10^{-5} , and 10^{-4} , then increases slowly again at last with P_c around 10^{-3} , 10^{-2} , and 10^{-1} .

3. Membership function of time to TCA

The times to TCA include the time from current to TCA (denoted as T_1) and the time from last observation to TCA (denoted as T_2). The two type of time are in negative numbers for the convenience to trend analysis. Similar to collision probability, the membership function of times to TCA is chosen to be curvilinear, which can be expressed as

$$f_T = \begin{cases} e^{-kT^4} & L_T \leq T \leq 0 \\ 0 & T < L_T \end{cases} \quad (6.96)$$

Let $k_1 = 0.005$ and $L_1 = -7$ days for the time from current to TCA (T_1); $k_2 = 0.00125$ and $L_2 = -10$ days for the time from last observation to TCA (T_2). The curve of member function is shown in Fig. 6.37.

4. Membership function of orbit quality

The quality of orbit determination is rated from 0 to 10 by experts considering the accuracy of observations and the orbit determination model. The linear member function of orbit quality is used to map 0–10 to 0–1, in which 0 means the worst and 1 means the best, as shown in Fig. 6.38.

Fig. 6.37 Membership function curve of times to TCA

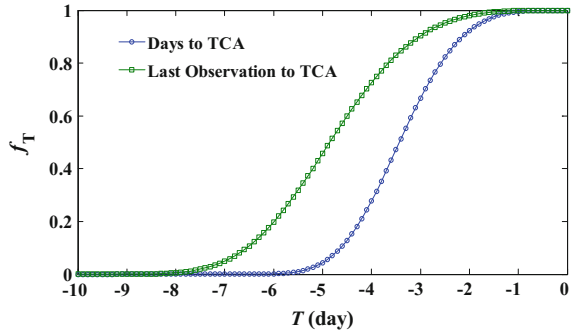
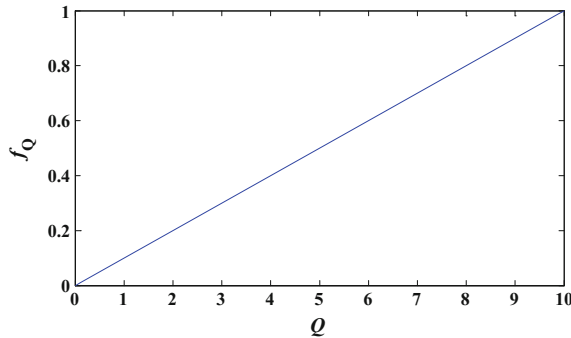


Fig. 6.38 Membership function curve of quality of orbit determination



$$f_Q = \frac{Q}{10}, \quad 0 \leq Q \leq 10 \tag{6.97}$$

6.4.4 Weighting Coefficients of Assessment Parameters

The analytic hierarchy process (AHP) evaluation method [19], in accordance with the performance evaluation system from the theory of system engineering, is chosen to calculate the weighing of each parameter. The method is the synthesis of quantitative and qualitative analysis, which is used to determine the relative importance of each parameter by way of pairwise comparisons, and establish the judgment matrix by calculating the maximum eigenvalue and eigenvector. The weighting vector ω of each parameter is just one of the values of maximum eigenvector. Following is the process steps:

(1) Construction of judgment matrix

The AHP method requires the decision-makers to give out the relative importance of each parameter scaled 1–9 in judgment matrix A

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix} \quad (6.98)$$

where a_{ij} indicates the relative importance of indicators u_i to u_j . Elements of matrix \mathbf{A} are given by experts, usually take 1–9 and their inverse number. That means:

If $a_{ij} = 1$, u_i is as important as u_j ;

If $a_{ij} = 3$, u_i is a little more important than u_j ;

If $a_{ij} = 5$, u_i is obviously more important than u_j ;

If $a_{ij} = 7$, u_i is strongly more important than u_j ;

If $a_{ij} = 9$, u_i is absolutely more important than u_j ;

The same meaning is given when the number is equal to 2, 4, 6, 8 and their inverse number.

(2) Weights calculation of each parameter

The weights of each parameters is calculated by solving $\mathbf{A}\boldsymbol{\omega} = \lambda_{\max}\boldsymbol{\omega}$, λ_{\max} denotes the maximum eigenvalue, $\boldsymbol{\omega}$ denotes the corresponding eigenvector, the components of $\boldsymbol{\omega}$ is the weights of each parameters.

(3) Consistency test

Since many factors are involved in the comprehensive evaluation, consistency test of judgment matrix should be taken to eliminate the inconsistency from different experts. The index of consistency test is defined as

$$CI = \frac{1}{m-1}(\lambda_{\max} - m) \quad (6.99)$$

The judgment matrix has full consistency if $CI = 0$. The higher the CI , the worse the consistency of judgment matrix. To test whether the judgment matrix satisfying consistency, the random consistency index RI is used to compare with CI . Define the random consistency ratio as $CR = CI/RI$. If $CR < 0.1$ the judgment matrix is considered to satisfy consistency, and the weights vector $\boldsymbol{\omega}$ is reasonable. Otherwise the judgment matrix should be reconstructed until the consistency is satisfied.

We consider the R miss distance has greater effect, the S and W miss distance has the same effect on the risk assessment of collision. The time from current to TCA and the time from last observation to TCA are equally important, and orbit quality is more important. The judgment matrices of risk assessment parameters and quality assessment parameters are constructed based on the consideration, as shown in Tables 6.19 and 6.20.

By calculating the weights vector, the weighting coefficient of each parameter can be obtained as shown in Table 6.21.

Table 6.19 Judgment matrix of risk assessment parameters

	P_c	R	S	W
P_c	1	1/3	2	2
	3	1	2	2
S	1/2	1/2	1	1
W	1/2	1/2	1	1

Table 6.20 Judgment matrix of quality assessment parameters

	T_1	T_2	Q
T_1	1	1	1/3
T_2	1	1	1/3
Q	3	3	1

Table 6.21 Weighting coefficient of each parameter

Risk assessment parameters				Quality assessment parameters		
P_c	R	S	W	T_1	T_2	Q
0.247	0.436	0.158	0.158	0.2	0.2	0.6

6.4.5 Examples

US-Russian satellite collision event is chosen to test the above method. Nine sets of pre-collision TLE data downloaded from space-track website have been used to confirm the actual collision event in space. In analysis, the orbit quality, the time from current to TCA, the time from last observation to TCA are used as quality assessment parameters. RSW miss distance at TCA and collision probability is used as risk assessment parameters. Table 6.22 shows the risk assessment parameters and quality assessment parameters of the nine data sets.

Table 6.22 Risk assessment parameters and quality assessment parameters of US-Russian satellite collision event

T_1 /day	T_2 /day	d_M /km	R /km	S /km	W /km	P_c	Q
-1.2072	-1.3397	0.5838	0.0403	0.3647	-0.4541	6.2×10^{-4}	9
-1.8978	-2.1858	0.8414	0.0497	0.5257	-0.6551	3.9×10^{-4}	8
-2.0374	-2.2557	0.9836	0.0446	0.6152	-0.7662	3.5×10^{-4}	7
-2.8048	-2.8150	0.8231	0.0697	0.5133	-0.6396	2.7×10^{-4}	9
-3.2234	-3.5839	0.6881	0.1016	0.4260	-0.5308	1.2×10^{-4}	8
-3.7936	-3.9211	0.9894	0.0761	0.6174	-0.7694	1.8×10^{-4}	7
-5.3165	-5.7509	0.1170	0.1127	-0.0196	0.0243	6.3×10^{-5}	6
-5.8165	-6.2402	0.1259	0.1246	0.0115	-0.0143	4.8×10^{-5}	7
-8.3165	-8.3373	0.7578	0.2056	0.4565	-0.5689	2.3×10^{-5}	8

Figures 6.39 and 6.40 give out the evolution curves of miss distance and collision probability. Figure 6.41 shows the evolution curves of f -values of quality assessment parameters. Figure 6.42 shows the evolution curves of F -value and f values of risk assessment parameters.

Fig. 6.39 Evolution curve of RSW miss distance

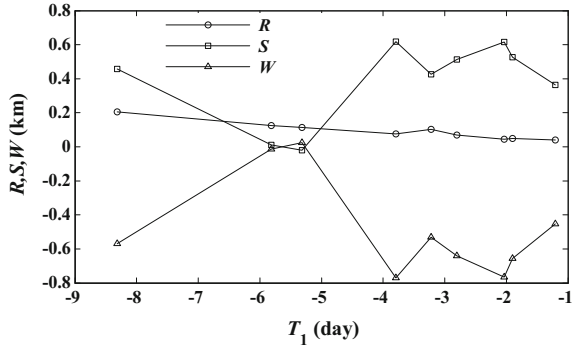


Fig. 6.40 Evolution curve of collision probability

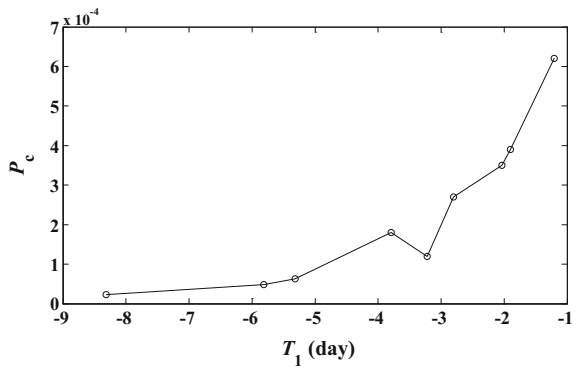


Fig. 6.41 Evolution curve of the f -value of quality assessment parameters

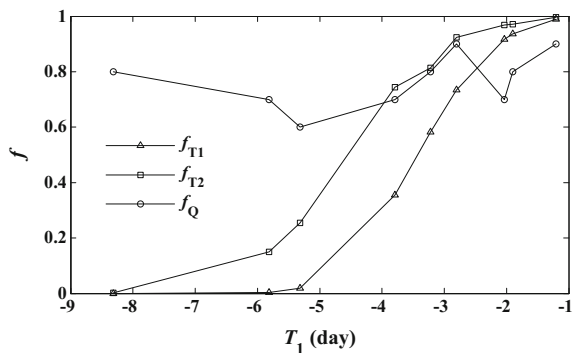
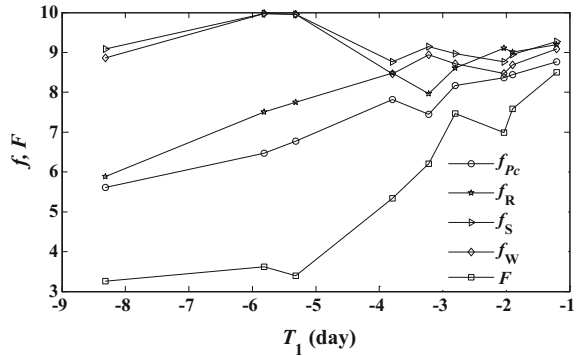


Fig. 6.42 Evolution curve of the f -value of risk assessment parameters and the F -value



These evolution curves are helpful to the comprehensive assessment of conjunction risk. Figure 6.42 indicates that the F -value is becoming larger when the time is closing to TCA (about 9), which means in the actual US-Russian satellite collision event, the risk of collision is steadily growing.

References

1. Newman LK, Duncan M (2006) Establishment and implementation of a close approach evaluation and avoidance process for Earth observing system missions. In: AIAA 2006-6291, AIAA/AAS astrodynamics specialist conference and exhibit, keystone, Colorado, 21–24 Aug 2006
2. Coppola VT, Woodburn J, Hujsak R (2004) Effects of cross correlated covariance on spacecraft collision probability. AAS 04-181. The AAS/AIAA space flight mechanics meeting
3. Jenkin AB (2004) Effect of orbit data quality on the feasibility of collision risk management. J Spacecraft Rockets 41(4):677–683
4. Frigm RC, Levi JA, Mantziaras DC (2010) Assessment, planning, and execution considerations for conjunction risk assessment and mitigation operations. In: AIAA 2010—1926. SpaceOps 2010 conference: delivering on the dream, Huntsville, Alabama, 25–30 Apr 2010
5. Alfriend KT, Akella MR, Frisbee J et al (1999) Probability of collision error analysis. Space Debris 1:21–35
6. Bai XZ, Chen L (2009) Explicit expression and influencing factor analysis of collision probability between space objects. Chin J Space Sci 29(4):422–431
7. Alfano S (2003) Relating position uncertainty to maximum conjunction probability. In: AAS 03—548. AAS/ AIAA astrodynamics specialists conference, Big Sky, Montana
8. Bai XZ (2008) Research on collision probability in space objects collision detection. National University of Defense Technology, Changsha
9. Alfano S (2004) Determining probability upper bounds for NEO close approaches. In: AIAA 2004—1478 planetary defense conference: protecting earth from asteroids, Orange County, California
10. Gottlieb RG, Sponaugle SJ, Gaylor DE (2001) Orbit determination accuracy requirements for collision avoidance. In: AAS 01—181. AAS/AIAA space flight mechanics meeting, Santa Barbara, California
11. Peterson GE (2004) NEO orbit uncertainties and their effect on risk assessment. In: AIAA 2004—1420. Planetary defense conference; protecting earth from asteroids, Orange County, California

12. Chan FK (2008) Spacecraft collision probability. The Aerospace Press, El Segundo
13. Klinkrad H (2006) Space debris-models and risk analysis. Springer-Praxis, New York
14. Alfano S (2014) Determining a probability-based distance threshold for conjunction screening. *J Spacecraft Rockets* 50(3):686–690
15. Laporte F (2014) JAC software, solving conjunction assessment issues. In: Proceedings of Advanced Maui Optical and Space Surveillance Technologies Conference, Maui, Hawaii, 2–3 Sep 2014
16. Frisbee JH (2015) An upper bound on orbital debris collision probability when only one object has position uncertainty information. In: AAS 15-717. AAS/AIAA Astrodynamics Specialist Conference, Vail, CO, 9–13 Aug 2015
17. Ma ZH (2000) Handbook of modern applied mathematics: volume for probability statistics and stochastic process. Tsinghua University Press, Beijing
18. Frigm RCA (2010) Single conjunction risk assessment metric: the F—value. *Adv Astronaut Sci* 135:1175–1192
19. Tan YJ, Chen YW, Yi JX (1999) Fundamentals of systems engineering. National University of Defense Technology Press, Changsha

Chapter 7

Orbital Anomaly and Space Events Analysis

7.1 Overview of Orbital Anomaly and Space Events

7.1.1 *Orbital Anomaly and Space Events*

Space events may be related to spacecraft maneuver, in-orbit collision, explosion and disintegration, revulsions in trajectory coefficients and space environment. All of them can give rise to the anomalies of orbits of space objects. Space events are mainly produced by two causes: one is artificial controls, like orbital maneuver of spacecraft; the other is non-artificial controls, such as collision, explosion, disintegration, and changes in space environment.

The detection of space events is an important content for space surveillance and space situation awareness. Prompt awareness of space events has become increasingly important for spacecraft owners and operators. All they need know is what and when happened to evaluate the risk or take appropriate measures.

There are two goals for space events detection [1]: First, near real-time detection of space events like new object appearance, object maneuver, and object disintegration are indispensable functions for any space situation awareness system. Second, it often needs to make a list of previous space events to support near real-time detection and related general analysis.

Among the orbital anomalies of non-artificial control, in-orbit collision, explosion, and disintegration will produce space junks that may collide with other moving spacecrafts. The abrupt change in space environment will have greater effects on space missions and orbital movements of spacecrafts. Hence, accurately identifying these space events can help analyze the potential causes of abnormalities and solve them with countermeasures.

For the orbital anomalies of artificial control, detecting and identifying maneuvers of spacecrafts, especially other parties' spacecrafts, are conducive to the accurate orbital tracking and determination of our side. Based on this, we can judge whether the spacecraft works normally, and at the same time, analyze the intention

and purpose of maneuver to determine whether it is dangerous to our side and whether the maneuvered orbit can set other spacecraft on collision risks. Moreover, unknown maneuver will complicate the close approach analysis and collision warning.

GEO is located in a specific zone with scarce resources, where is full of space objects. Affected by perturbations, GEO satellites need frequent position-holding maneuvers. It is thus especially important to detect orbital anomalies for GEO objects.

1. Analysis of spacecraft working conditions based on orbital anomaly detection

There is only a few orbital spacecrafts that are always in working conditions; most are malfunctioned. As space forces exert greater impacts on global politics, economy, and military affairs, to evaluate the working conditions of foreign spacecrafts has become a vital task for space surveillance. A complete space event detection system can be used to analyze the time and nature of the space event, so that we can know what measures are needed to be taken. To evaluate working conditions and anomalies of foreign spacecrafts is a vital task for Space Situation Awareness.

The spacecrafts in normal working conditions are with controllable orbits. The following orbit control activities have to be conducted to finish corresponding tasks: orbital maneuver, orbit keeping, attitude adjustment, rendezvous and docking, and so on. And hence, relatively drastic changes will occur to orbital elements of spacecrafts. Spacecrafts in abnormal working conditions are with uncontrollable orbits; only under the effects of gravity and different preservative forces, can they fly without restraint along their orbits. Except in the cases of in-orbit collision, explosion, and disintegration, their orbital parameters (such as orbital elements and other derived parameters) will vary smoothly and continuously. Thus, the working conditions of spacecrafts can be determined according to differences in orbital characteristics.

2. Analysis of unusual space weather based on orbital anomaly

Space situation anomalies involve space weather anomalies and orbital anomalies of space objects. Unusual space weather can affect the satellites in orbit. To identify spacecraft maneuvers with military intentions in good time can save the warning and response time of our side, and detections of in-orbit collisions and disintegrations are convenient for corresponding avoidance measures.

Space weather factors with greater impacts on orbits comprise solar activities and earth magnetic fields. Solar activities will add more high-energy particles in the upper atmosphere and increase their energies to increase the temperature of the upper atmosphere. Coronal mass ejections can influence earth magnetic fields to cause geomagnetic activities and heat up the upper atmosphere. The rising of upper atmosphere temperature will further lead to volumetric swelling and enhance the atmosphere density. As a result, the atmospheric drag against LEO objects will be increased to influence the orbit energy attenuation rate of these objects and enlarge

the uncertainty in orbit prediction. The changes brought by these activities to the upper atmosphere density will directly influence spacecraft orbits and make the attenuation rate of orbital energy have an abrupt change. To this end, we can study the relevance between anomalies of spacecraft orbits and space weather events after knowing in what way the spacecraft orbits are affected by the changes in solar and geomagnetic activities.

7.1.2 Orbital Anomaly Detection Based on Historical Data

Historical orbital data for space objects are often used to detect space events. For inactive objects like spacecraft with detailed information difficult to obtain, rocket bodies and space debris, to detect space events with orbital elements is a feasible alternative solution. This requires a dynamic orbital database. The Space Situation Awareness system can provide previous and current orbital elements for space object. The previous orbital data is an array of time series of orbital elements. Through analyzing these series we can get to know the changing situations of orbital elements and discover orbital anomalies. It is done based on Space Situation Awareness information and a universal space situation map, because prior and real-time orbital elements and corresponding epoch time need to be obtained from the space situation information.

The aforesaid detection methods can be divided into two classes: One is statistical methods, such as the Patera approach [2] and Kelecy approach [3]. It requires the statistical analysis of the time series of orbital elements of a certain object to find out when the orbital elements will become exceptional ones relative to historical information. The other is the Hujsak approach [4] and the like, namely, using the prediction of orbital elements or state vectors to determine if subsequent ephemerides is the natural propagation of previous ephemerides [5]. For GEO, special detection methods have been put forward for detection of satellite maneuvers based on evolutions of orbital characteristics, like Flohrer approach [6].

Because of various reasons, there are always abnormal values found in TLE time series of orbit elements of space objects. While the orbital elements change with time under the action of disturbing forces, a few will deviate significantly from the normal state. It may be resulted from many reasons. In statistics, outliers can be caused by multiple factors, for instance, experimental conditions varying, unknown new phenomena cropping up, system errors, or extreme values produced on account of random error-induced extreme volatility in estimated values. In TLE time series of orbit elements, there are two primary causes for outliers: sudden change of orbital elements caused by space events and gross residual errors in TLE fitting.

The analysis of orbital anomaly based on orbital elements is a basic approach for space event detection. Using the variation characteristics of selected orbital elements to detect orbital anomalies is based on data fitting, outlier detection, and other data processing algorithms. The methods adopted at present include Moving Window Curve Fitting, Binary Search, and Extended Kalman Filter, etc.

Some scholars at home and abroad have analyzed the detection methods for space events or orbital anomalies using historical data. Patera from Aerospace Corporation of U.S. put forward a detection method for space events based on Moving Window Curve Fitting (MWCF) [2]. This method is related to the polynomial fitting of data in a window with a certain length. The fitted values were subtracted from cataloged values to get the dispersion values and the window was moved to obtain the data sequences of the dispersions. The detection threshold was set based on standard deviations of the dispersion data. A space event occurs when dispersions between the parameters and their expectation values exceeded the detection threshold. This method is computationally efficient and does not require derivation and prediction of orbital state.

Kelecy et al. at the American Boeing Co. proposed a method in detecting satellite maneuvers with historical TLE data and pointed out its limitations [3, 6]. This method is analogous to what Patera applied, that is, selecting two adjacent data segments of the same certain length, conducting polynomial fitting and smoothing in these two segments, and then calculating the difference of predicted values in the center points of the two propagation periods. The data segments are moved to obtain dispersion data sequences and the outlier detection method is employed to determine the anomalous changes of orbital parameters. Kelecy et al. as well used the radar observation data of space objects to study the detection method for low thruster maneuver with magnitude of cm/s speed increment [6].

Swartz and his colleagues from the American Aerospace Corporation proposed a method to detect space events based on Space Incident Flagging Technique (SIFT) [7], which is a method used to monitor the satellite ephemerides to detect changes that are out of the ordinary compared to the object's historical behavior based on the method Patera once used. This improved method is divided into two steps: first, put intermediate results, i.e., dispersions (or derivatives), into the database; second, shorten the computation time through storing these intermediate results in the database. SIFT has also been applied in the Space Situation Monitoring Laboratory [8], as described in the document [7]. The problems in applications have been discussed and suggestions for improvement and correction have been given to detect satellite anomalies on time.

Europe [9, 10] and the United States [11, 12] have established their own application systems for Space Situation Awareness to detect orbital anomalies of space objects. The aforementioned methods have been applied in practice.

In China, Dong Yunfeng et al. at the Beihang University have come up with a method to identify and detect orbital maneuvers of space objects based on wavelet analysis [13, 14]. Orbital maneuvers are associated with mechanical energy changes. Mechanical energy cannot be directly measured; it can only be calculated by the position and velocity of a space object instead. The range and angle information of a space object can be measured by ground-based radar and then transformed into positional information. The velocity too can only be obtained through differential smoothing based on positional information. The changes induced by maneuvers to the mechanical energy are almost unidentifiable on account of measurement noises. Wavelet Transform (WT), a algorithm with multi-resolution capability, can be used

to reveal the mechanical energy change that has been drowned in noise to enhance the recognition capability of orbital maneuvers. Binary wavelets were used to analyze the mechanical energy changing with time under different scales. True signals were found increasing in amount and noise levels suppressed as the scale grew larger. The wavelet coefficient curve varying with the wavelet analysis scale was observed to determine whether there were orbital maneuvers or not. This method only requires measurements of space object distances and angles using ground stations to identify orbital maneuvers.

Qiang Sheng from the National University of Defense Technology has provided a fusion algorithm to identify the orbital maneuver of space object through a spaced-based space surveillance system [15]. In addition to searching, capturing, tracing, data processing, and orbit determination, whether the orbit can be retraced or not after orbital maneuver is an essential function for a space surveillance system. For non-cooperative objects, it has an important realistic significance to know the exact time and variation principles of the orbital maneuver in time. The specific intention of orbital maneuver can be obtained through studying and analysis, which is crucial for change of space surveillance from passive mode to active mode.

Yang Xu et al., from the National Space Science Center, CAS, developed a comprehensive criterion method [16] to analyze the orbital anomalies of LEO satellites based on TLE data. He pointed out that orbital anomalies were caused mainly by two reasons: one is artificially control of orbital maneuvers; the other is non-artificial control of orbital changes, for instance, the orbital changes resulted from space debris bumping with satellites or fuel leak. To discover the orbit anomaly of an in-orbit satellite timely and accurately is of great significance, especially for non-artificial control of orbital anomalies. Possible influencing factors can be analyzed immediately after anomalies are found. Yang Xu's colleagues Wang Ronglan et al. also proposed a simple and efficient method to detect orbital anomalies based on semi-major axis changes [17, 18].

All the methods introduced above belong to MWCF, which is about processing the data based on mathematics without considerations of orbital models or predicting characteristics. These methods can be applied to all types of orbits. The GEO area is a key area for space surveillance and space situation awareness, where large amounts of space orbits are gathered [19].

Aaron [20] and Folcik [21] presented three ways to detect GEO object maneuvers based on intensive observation data: Binary Search Algorithm, Adaptive Extended Semi-Analytical Kalman Filter, and Differential Correction-Extended Semi-Analytical Kalman Filter Hybrid Algorithm. The Hybrid Algorithm is the best of them. However, none of the three methods are suitable for TLE data-based orbital anomaly detection, because the TLE data is too sparse to conduct differential correction or filtering.

Abbot and Wallace in the MIT Lincoln Laboratory have introduced a method to detect GEO satellite maneuvers based on the data of satellite orbits and property information [22]. Satellite property information includes image information of optical telescopes and RCS information of radars. Automatic information fusion technology was used to combine the property information with the orbit

information and put the data into a Bayes Belief Network (BBN) to detect the state changes of GEO satellite and monitor the situations of non-cooperative GEO satellites.

7.1.3 Basic Procedure of Detecting Orbit Anomaly

In this chapter, the orbital anomalies of space objects are detected based on the viewpoint that system anomaly is that the actual value of some parameter of the system deviates from its desired value to an incredible extent. The desired value is generally reckoned from a system model or extrapolated from historical data instead of system model prediction and historical data fitting propagation. The actual value is generally obtained by observation or conversion from the observed value. Thus errors are unavoidable. It is related to the elimination of noise and abnormal values.

If the viewpoint explained above is used to detect orbital anomalies of space objects, the orbital anomalies would be that the actual values of some characteristic orbit parameters of the orbit deviate greatly from the desired values thereof. Based on this theory, the detection method should include the following procedures:

- (1) Selection of characteristic orbit parameters. The characteristic orbit parameters that can fully reflect the orbital features of space objects should be determined above all. Different parameters reflect different space events or orbital anomalies. The characteristic parameter is one or multiple orbital elements or derived parameters. Its changes can adequately mirror the orbital anomalies of space objects and can be used to analyze the causes of orbital anomalies. Different types of orbits should have different characteristic orbit parameters.
- (2) Generation of Dispersion between actual and expectant value. The generation of dispersion requires sufficient precision in actual and expectant value. The expectant value is obtained from an orbital model or historical data while the actual value is observed. In this chapter, the actual value is calculated with TLE data. Comparing it with the expectant value can get the dispersion of characteristic orbit parameters.
- (3) Anomaly detection of dispersion value. Because of the errors in models and data, the actual value always differs from the expectant value regardless of orbital anomalies. So, after obtaining the dispersion data, some criteria are needed to determine if the dispersion is normal. This is a matter of detecting data anomalies. Since more than one characteristic parameter might be selected, the dispersion data of several characteristic orbit parameters must be considered at the same time in anomaly detection, and the components of dispersion data need be uniformly detected in a multidimensional way.
- (4) Evaluation of anomaly detection. Whether the detection of orbital anomalies is reliable depends on precisions of the actual and expectant value and configuration parameters of the detection method. In fact, orbital anomaly detection is a matter of identification and analysis. False alarm and missing alarm are

always inevitable. The calculation and estimation of misjudgment probability is also indispensable during anomaly identification.

- (5) Analysis of anomaly causes. The causes should be analyzed after orbital anomalies are discovered. Different causes lead to different orbital anomalies. According to various reasons, there are always a few abnormal data in the time series of orbital elements. Abnormal data should also be analyzed.

According to the thinking above, we have introduced in this chapter three ways to detect the orbital anomalies of LEO and GEO objects based on previous orbital data sequences: MWCF, Secular Orbital Dispersion (SOD), and Drift Rate Dispersion (DRD). These three methods are universal in their own formats of object catalogs. Theoretically, the MWCF and SOD methods are applied to all orbits, while the DRD method is suitable for GEO only. The difference lies in the selection of characteristic orbit parameters and the generation method of dispersion data. See the basic procedures of the three methods in Fig. 7.1.

First, the selection of characteristic orbit parameters needs to consider the features of different orbits and the orbital maneuvers of spacecraft. As space events (maneuver, collision, explosion, sudden change of space weather, etc.) chiefly alter orbit energy and inclination angles, total orbit energy (the sum of kinetic and potential energy) and orbit inclination angles have been selected as characteristic orbit parameters for the MWCF method. The semi-major axis and inclination angle of the object have been selected as the characteristic orbit parameters for the secular-term dispersion method. In consideration of the satellite position and inclination angle drifting of GEO, the mean longitude of the satellite and orbit

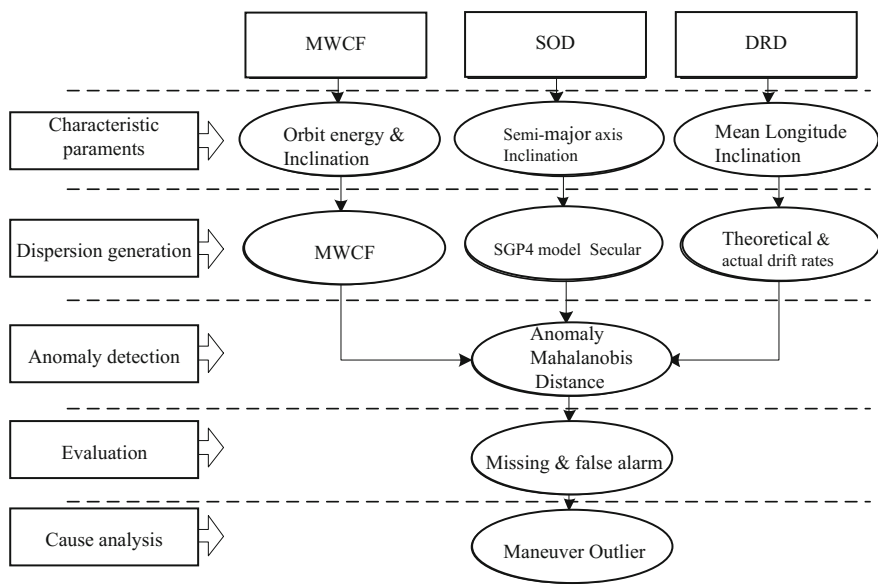


Fig. 7.1 Basic procedure of detecting orbit anomaly

inclination angle have been selected as characteristic orbit parameters for GEO drift rate dispersion. The dispersion data have been calculated with the drift rates of mean longitude and inclination angle.

Second, the generation method of dispersion data is virtually a calculation method of theoretical and actual parameters.

The theoretical parameter value in MWCF is obtained through polynomial fitting in the data window, and the actual parameter value is selected to be the calculation value of the TLE data of the last point in the data window, and then the dispersion data of orbital energy and inclination angle are obtained by the fitting value with the actual value at the current time. Regarding the SOD method, the theoretical parameter value is propagated by the TLE at one or a few previous epoch times with the SGP4 model which just considers the secular terms. The actual parameter value is the TLE calculation value at the current time. The difference between the actual and theoretical value is divided by the epoch difference to get the change rate of orbital parameters, and this change rate is used as dispersion data. Regarding the GEO drift rate dispersion method, the characteristic orbit parameters are drift rates. The theoretical drift rate is calculated with the orbital elements at the current time, while the actual one is obtained by fitting or smoothing with historical data. The difference between the two is used as dispersion data.

The dispersion data between the actual and theoretical parameter values are hence obtained through the two procedures above. There is more than one characteristic parameter, so the dispersion data is a multidimensional vector with respect to time. A criterion should be adopted to determine whether it is in a normal status or not. Several components of the dispersion data are uniformly checked with the detection method of multidimensional data anomalies. In this chapter, the Mahalanobis Distance is adopted as a measurement for the distance from each sample to its distribution center. If the Mahalanobis Distance is larger than the preset threshold value, the dispersion data is considered abnormal and orbital anomalies may exist at the corresponding time.

7.2 Orbital Anomaly Detection Based on MWCF

In this section we use Patera's MWCF method [2] to detect the orbital anomaly with our selection of characteristic orbit elements and Mahalanobis-distance-based outlier detection method.

7.2.1 Selection of Characteristic Orbit Elements

The selection of characteristic orbit elements is the first thing to be solved in the anomaly analysis of space object orbits. The characteristic orbit element is one or more orbit elements or derived parameters. Orbital anomalies can be fully reflected

and their causes analyzed by variations in these parameters. The selection of appropriate characteristic orbit elements is the foundation for orbital anomaly detection.

Orbital anomalies are generally divided into two classes: One is sudden changes to the orbital shapes and sizes in an orbital plane, which is chiefly embodied in the changes in semi-major axis and eccentricity; the other is sudden changes to the spatial position of an orbital plane, which is primarily reflected in the changes in orbit inclination angle and right ascension of ascending node (RAAN). In this book, the two classes are called coplanar anomaly and hetero-planar anomaly, respectively.

Space events refer to abnormal changes in orbital elements. There are multiple types of orbital parameters to detect space events. In consideration of the coplanar and hetero-planar anomalies, the parameters that can be used to detect orbital anomalies comprise orbit semi-major axis a , orbit energy E , eccentricity e , orbit inclination angle i and angular momentum h , etc. There are two reasons for not selecting the RAAN Ω and argument of perigee ω . First, Ω and ω are undefined for small inclination angle and near-circular orbits. It is inconvenient to use them to describe orbital changes. Second, under the influence of perturbative forces, Ω and ω vary rapidly with time; orbital anomalies have less effect on their variations compared to their drift distance. So, it is inconvenient to detect orbital anomalies.

Figure 7.2 shows changing curves of a , i , E , and h of an American satellite Terra.

As shown in Fig. 7.2, the varying patterns of a , i , E , and h are the same. Therefore, a (or E) and i are selected as characteristic orbit parameters to detect LEO maneuvers.

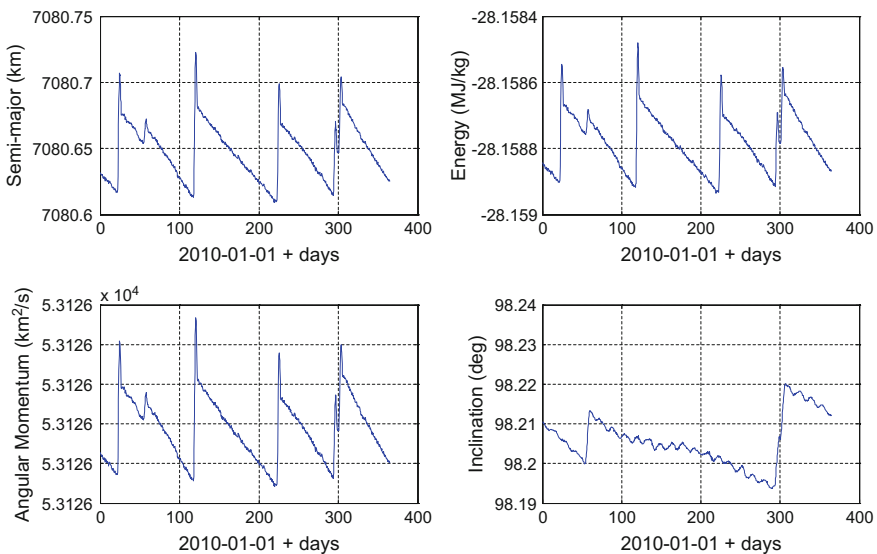


Fig. 7.2 Curve of semi-major axis, energy, angular momentum, and inclination versus time (Satellite Terra)

7.2.2 Generation of Dispersion Data

The MWCF technology is efficient in estimating parameter value and detecting parameter value mutation. It can be used to eliminate minor parameter changes caused by orbital perturbations and noise so as to extract secular variations and at the same time to filter the noise and deal with time varying data. Space events will occur when the dispersion between parameters and their expectant value surpasses the preset threshold value. With this method, the data of space orbits can be analyzed and the events, such as spacecraft maneuvers and collisions and sudden variations of space weather, can be easily detected. This method is computationally efficient and does not require the derivation and prediction of situation vectors.

The MWCF method is similar to the moving average method in random data processing. For each processing procedure, the single typical parameter value that had been obtained based on the polynomial fitting was subtracted from the actual parameter value with corresponding time to get the dispersion. The parameter dispersion and corresponding time were stored. The stored parameter dispersion was statistically treated to determine the mean value, variance, and standard deviation. The event threshold was designated by the standard deviation. All the dispersions were compared to see whether they were higher than the threshold or not. If so, it would indicate space events existed. At the same, the time and dispersions were stored for further analysis. Analyzers can adjust the threshold to detect most events that interest them.

Also, there is one other way. The first step was to obtain the polynomial fitting derivatives on the midpoint or other points of a data window. The second step was to substitute the derivative dispersion of parameters for parameter dispersion so as to solve false alarm that had been caused by slowly accumulated perturbations and long time intervals among some TLE data.

Figure 7.3 shows the way in which data are respectively selected from four data windows via the MWCF method.

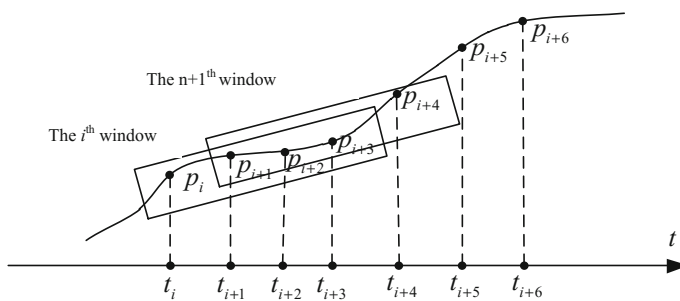


Fig. 7.3 Moving window curve fitting

After the moving window is selected, a polynomial for given orders was fitted with Least Squares (LS). The number of the order is generally set as three. See the equation below:

$$p(t) = c_3t^3 + c_2t^2 + c_1t + c_0 \quad (7.1)$$

Curve fitting have two effects. First, it can filter the high frequency variations of noise and data; second, it can calculate the parameter value of any time and its derivative value in the data window. Enlarging the data window can improve the noise filtering capability, but the size of large data windows needs fitting via a high-order polynomial curve to get the desired parametric variation. To this end, a balance must be maintained between window sizes and polynomial orders. The essential properties of window data can be employed to achieve this purpose. The output of curve fitting is expressed in polynomial coefficients c_i , as shown in Eq. (7.1). It should be noted that curve fitting is valid only in windows and cannot be applied to propagate the window.

Another option for detecting parameter variation is calculating the derivative value of fitting polynomial on several points. If step change of parameter was found in the center point of the window, the derivative of the fitting polynomial would see a wider change at this point. Then, the derivative value could substitute for the dispersion value obtained in the method above, with other procedures remaining the same. The expression of the polynomial derivative is listed below:

$$\frac{dp(t)}{dt} = 3c_3t^2 + 2c_2t + c_1 \quad (7.2)$$

The processing flow for generating data dispersion based on moving window polynomial fitting is shown in Fig. 7.4.

Specific steps are as follows:

- (1) Incorporate the followings into the continuous dynamic database: the code, epoch time, and state vector parameters (i , Ω , e , ω , mean anomaly M and angular velocity of mean motion n) of each object.
- (2) Select historical series of orbit elements for a space object.
- (3) Calculate the total energy E and inclination angle i that correspond to each array of orbit elements in the series based on the angular speed of the mean motion.
- (4) Take the subseries in the element series as moving window data. Here, the array of moving window data comprises time t , orbit energy E and orbit inclination angle i .
- (5) Do polynomial fittings of E and i in the moving window versus t and calculate the expectant values of E and i at the last epoch time in the window via polynomial fitting.
- (6) Subtract the expectant parameter value from the actual parameter value at the last epoch time in the window to obtain the dispersion data, and then store the epoch time and dispersion in the data array.

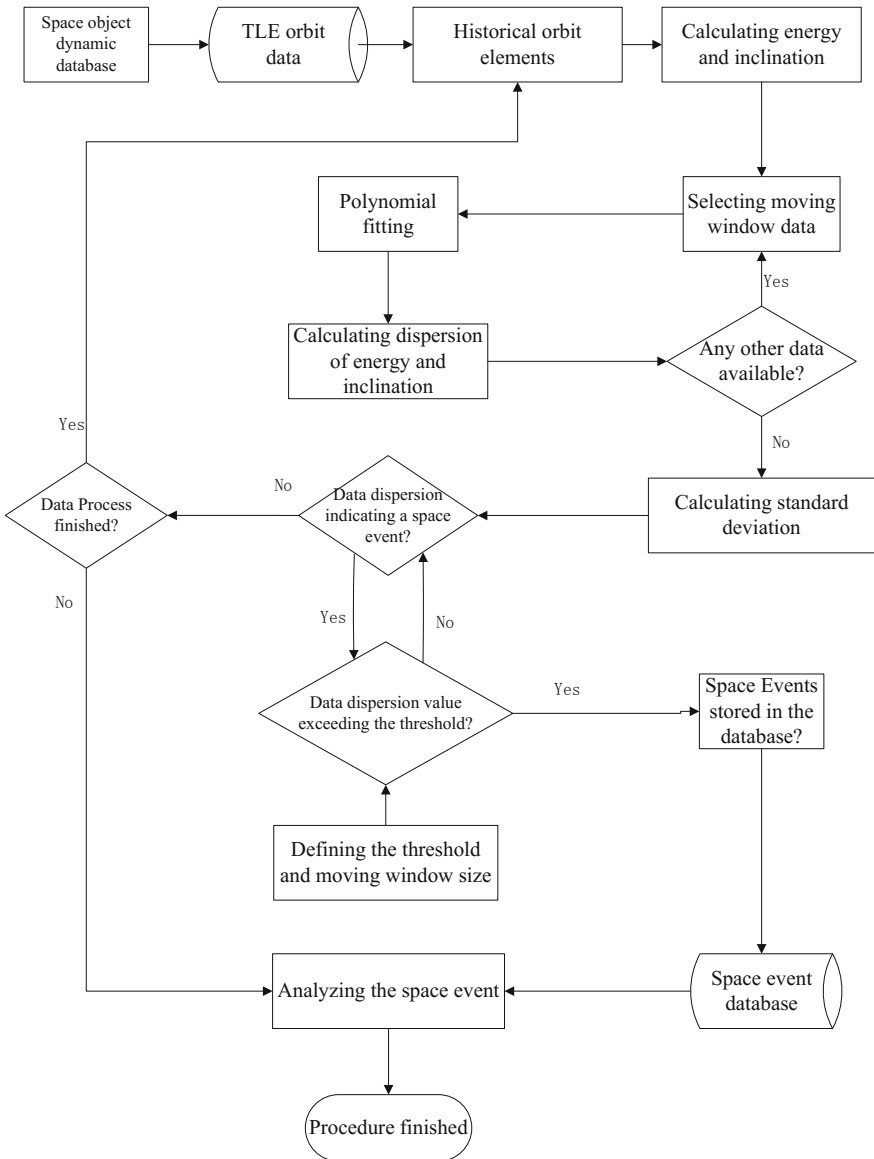


Fig. 7.4 Flow chart of moving window curve fitting

- (7) Judge whether there are any other data available. If so, update the moving data window, add the next data point, and delete the first one. At this time, the window remains the same in size, it just moving forward with time.
- (8) If there are no data available, calculate the standard deviation of dispersion data.

- (9) Judge whether each dispersion value represents a space event and check if it can be used for the detection of space events.
- (10) A space event occurs if the dispersion exceeds the result obtained by multiplying the standard deviation with the preset threshold. For instance, if the threshold is 3, then the standard deviation threshold is 3σ .
- (11) Store the space event in the database for the convenience of future analysis.
- (12) Check if there are objects left in the database for detection. If so, select the next space object and its data series to detect other space events. If not, analyze the existing space events.
- (13) Analyze each space event in the database.
- (14) Adjust the user-defined threshold to select the size of detecting events.

7.2.3 Detection of Outliers in Dispersion Data

Essentially, the analysis of orbit anomalies is based on the determination of the calculated actual and expectant values of characteristic parameters. If the dispersion is incredibly large, it is supposed to be abnormal. Orbit anomalies can be detected with the calculated dispersion data of orbit energy, semi-major axis, and inclination angle.

The dispersion data at each time can be obtained by subtracting the theoretical orbit energy, semi-major axis, and inclination angle from the actual ones. Further calculating the above change rates at each time point can get the data array of $\{t^{(k)}, \Delta D_E^{(k)}, \Delta D_a^{(k)}, \Delta D_i^{(k)}\}$, $k = 1, 2, \dots, n$.

These dispersion data were used to detect orbital anomalies. It is worth noticing that this method must be used together with the dispersion data of two characteristic orbit parameters, i.e., the semi-major axis a and inclination angle i .

The outliers in two components of dispersion data were detected both by multidimensional data and Mahalanobis Distance. The only information that could be obtained was data points; data variance and other information were unknown. At this time, outlier detection was like a learning approach without monitoring clusters [23].

A 2D vector set $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n\}$ has been established, with $\mathbf{r}_i = (x_i, y_i)^T$. The sample mean vector and covariance matrix are as follows:

$$\bar{\mathbf{r}} = \sum_{i=1}^n \mathbf{r}_i, \quad \mathbf{C} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{r}_i - \bar{\mathbf{r}})(\mathbf{r}_i - \bar{\mathbf{r}})^T \quad (7.3)$$

In outlier detection, the difference between two data points should be determined and the data points should be classified. This difference should be measured by a numerical value. To this end, a mathematical concept “distance” was abstracted to describe the difference between data samples. The smaller the distance was, the more closely the two samples became. The most common distances are Euclidean, Mahalanobis, B-mode, and Minkowski distances.

Euclidean distance is a well-known geometrical distance between two points in space. Useful as it is, it also has a few obvious defects. When the measurement unit is changed for instance, the distance value calculated is different. Moreover, when the components of the quantity index represent the substances of different quality or the components differ considerably from each other (for example, the deviation magnitudes of two change rates differ greatly from each other), large number annihilating small number would happen to the Euclidean distance. The distance needs weighing, and one of the best ways is to weaken the incidence of the components with large dispersion in distance. Mahalanobis, therefore, proposed the Mahalanobis Distance in 1936 to solve this problem [24, 25].

The population mean and covariance matrix were respectively set as $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma} > 0$, and the sample vectors of population were indicated by \mathbf{r}_i and \mathbf{r}_j . The Mahalanobis Distances from \mathbf{r}_i to \mathbf{r}_j and from \mathbf{r}_i to population were defined as below:

$$\begin{cases} d_M(\mathbf{r}_i, \mathbf{r}_j) \triangleq \sqrt{(\mathbf{r}_i - \mathbf{r}_j)^T \boldsymbol{\Sigma}^{-1} (\mathbf{r}_i - \mathbf{r}_j)} \\ d_M(\mathbf{r}_i, \boldsymbol{\mu}) \triangleq \sqrt{(\mathbf{r}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{r}_i - \boldsymbol{\mu})} \end{cases} \quad (7.4)$$

Mahalanobis Distance is dimensionless and unrelated to the measurement unit. What's more, it remains unchanged to all nonsingular linear transformations and immune to the selection of characteristic dimensions. According to the probability density function $\mathbf{r}: N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$,

$$\begin{aligned} f(\mathbf{r}) &= \frac{1}{(2\pi)^{\frac{3}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp \left[-\frac{1}{2} (\mathbf{r} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{r} - \boldsymbol{\mu}) \right] \\ &= \frac{1}{(2\pi)^{\frac{3}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp \left[-\frac{d_M^2(\mathbf{r}, \boldsymbol{\mu})}{2} \right] \end{aligned} \quad (7.5)$$

Mahalanobis Distance, actually, has described the eccentricity from the probability distribution center. The greater the Mahalanobis Distance from the sample value to the population is, the lower the probability in producing such sample becomes, vice versa. Because of this, Mahalanobis Distance was selected to measure the outlier degree of data points.

As the population mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ were both unknown during the time when the outliers in dispersion data of change rates were analyzed. Therefore, the unbiased estimations of sample mean $\bar{\mathbf{r}}$ and sample covariance matrix \mathbf{C} were used. The Mahalanobis Distance between data points \mathbf{r}_i and \mathbf{r}_j were as follows:

$$d_M(\mathbf{r}_i, \mathbf{r}_j) = \sqrt{(\mathbf{r}_i - \mathbf{r}_j)^T \mathbf{C}^{-1} (\mathbf{r}_i - \mathbf{r}_j)} \quad (7.6)$$

The Mahalanobis Distance between data point r_i and the sample mean were as follows:

$$d_M^{(i)} = d_M(r_i, \bar{r}) = \sqrt{(r_i - \bar{r})^T C^{-1} (r_i - \bar{r})} \tag{7.7}$$

In 2D outlier detection, Mahalanobis Distance was a measurement for describing the proximity of two data. The Mahalanobis Distance between the vector of a data point and the mean vector are used to show the outlier degree of this point. If it was higher than the given threshold d_T , this data point should be deemed as an outlier and removed.

The threshold d_T can be selected according to detection precision. The larger the threshold was, the lower the false alarm probability became, while the higher the missing alarm probability became, vice versa. Generally, d_T can be taken between 3 and 5. In this section, the threshold was set as $d_T = 5.0$.

Figures 7.5 and 7.6 show the Mahalanobis Distance curves for the dispersion data of satellites Terra and ERS-2. As can be seen in the figures below, the Mahalanobis Distances between several points are beyond 5.0. These data have been marked as outliers.

This method can also be explained with a 2D diagram. In the 2D space, the curves of equal Mahalanobis Distances ($d_M = 1, 3, 5$) are a series of ellipses as drawn in Figs. 7.7 and 7.8. If the threshold value is $d_T = 5$, the data points outside the ellipses of equal Mahalanobis Distances would be regarded as outlier points, which have been marked in red asterisks in the diagram below.

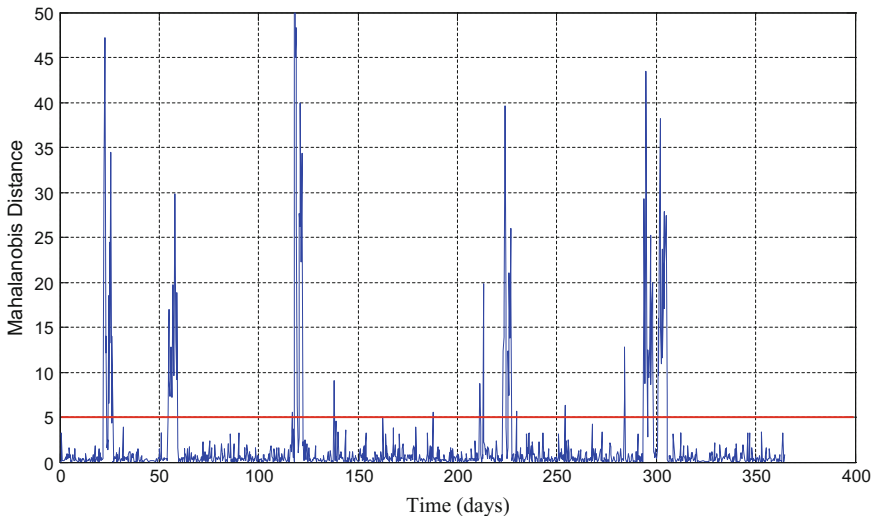


Fig. 7.5 Mahalanobis distance between the dispersion data and distribution center of satellite Terra

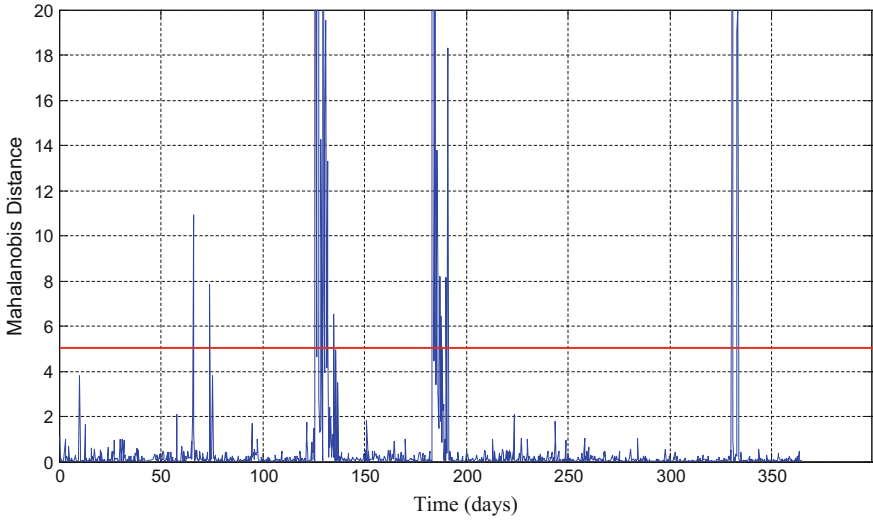


Fig. 7.6 Mahalanobis distance between the dispersion data and distribution center of satellite ERS-2

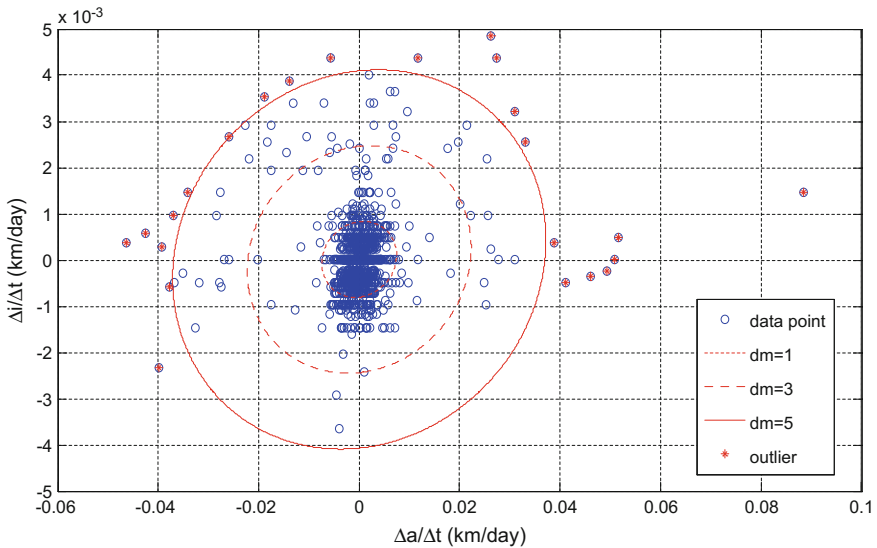


Fig. 7.7 Detection results of orbit outliers based on Mahalanobis distance (Terra)

The above contents have described the basic detection method for LEO satellite outliers based on semi-major axis and inclination angle dispersion. Figures 7.9 and 7.10 show the detection results of orbit outliers for satellites Terra and ERS-2. The parts circled in red are orbit anomalies.

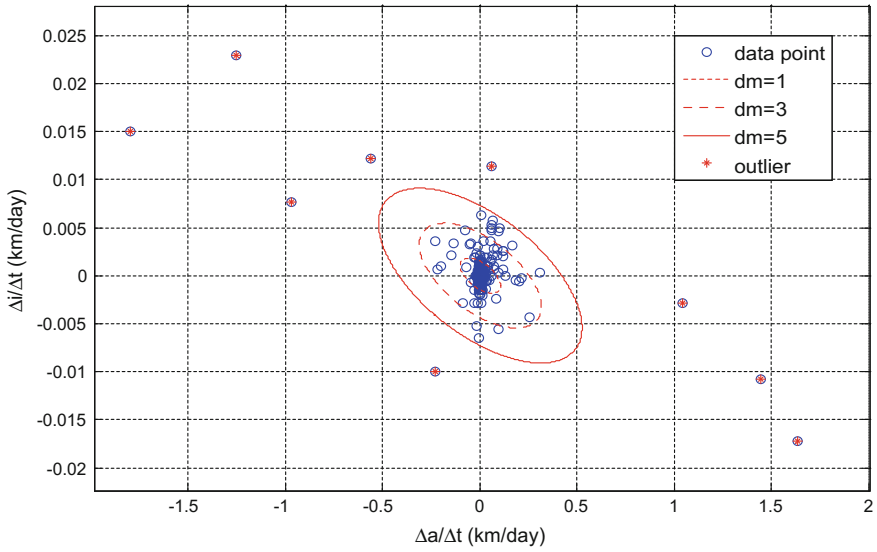


Fig. 7.8 Detection results of orbit outliers based on Mahalanobis distance (ERS-2)

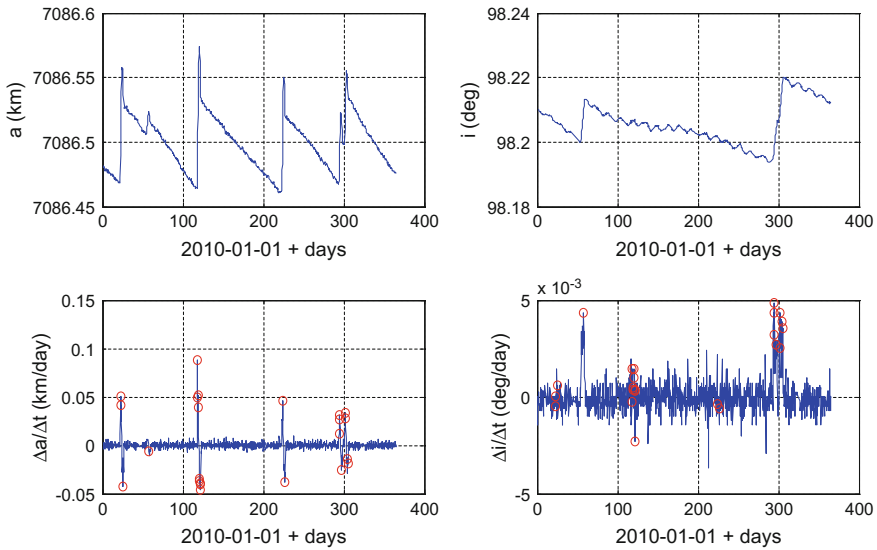


Fig. 7.9 Detection of Terra orbit anomalies in 2010

It's important to note that the orbital data points analyzed in this book are discrete. Anomalies may be all detected at the time points before and after the actual orbital maneuvering time, but the specific time and size of the orbital maneuver

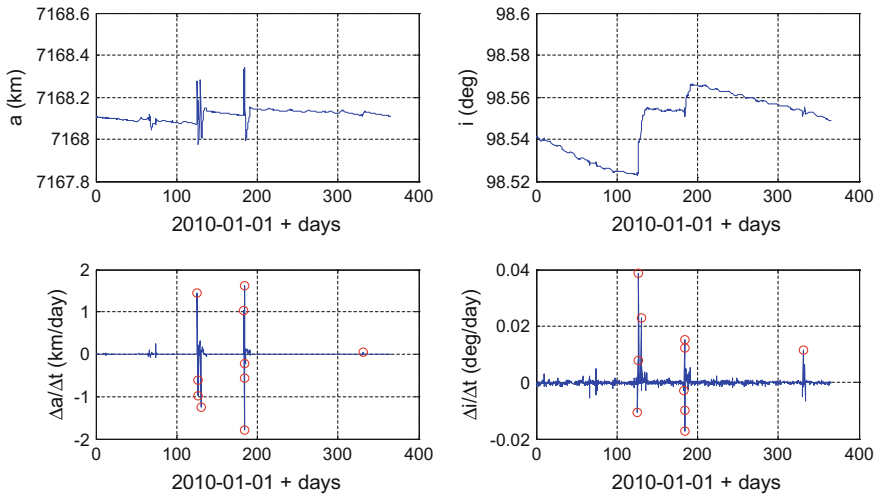


Fig. 7.10 Detection of ERS-2 orbit anomalies in 2010

need to be determined by a maneuver reconstruction method. What matters most is not the concrete time and size of the orbital maneuver, but whether there exist satellite maneuvers or not, which can be used to judge the working status of the satellite and facilitate the space object monitoring. So, in this book, the problem of maneuvering reconstruction is omitted.

7.3 Orbital Anomaly Detection Based on Prediction Dispersion

The MWCF-based detection method starts basically from data processing. Using a mathematical method to directly analyze the orbital data is unfavorable for studying the physical principles of orbital variations of space objects. In this section, both the orbit prediction and the mathematical detection method were combined and applied to the study of orbital anomalies.

TLE is the most complete catalog data of space objects released by NORAD at present. SGP4/SDP4 orbital prediction models are its analytical models. As an illustration, the historical TLE data has been analyzed in this section, and the secular term of the SGP4 model has been used to predict the secular variations of orbital elements. This method also applies to other types of orbital data and the prediction models thereof.

7.3.1 Secular Model of Characteristic Orbit Elements

To characterize the anomalies inside and outside the orbital plane fully, the semi-major axis a and inclination angle i have been selected as characteristic orbit elements in this section.

SGP4 is an orbit propagation model under the comprehensive influences of Earth nonspherical perturbation, atmospheric drag perturbation, and the third body attraction perturbation, and can be used for orbit prediction together with the NORAD released TLE.

TLE data include orbital elements and other relevant information of space objects, such as orbit codes, international codes, and atmospheric drag coefficients. These for SGP4 model prediction cover the followings:

$$\sigma_0 = (t_0, n_0, e_0, i_0, \omega_0, \Omega_0, M_0, B^*)^T \quad (7.8)$$

where t_0 is the epoch time; n_0 is the angular speed of mean motion (revolutions per day); e_0 is the eccentricity; i_0 is the mean orbit inclination angle ($^\circ$); ω_0 is mean perigee argument ($^\circ$); Ω_0 is the mean RAAN ($^\circ$); M_0 is mean anomaly at the epoch time ($^\circ$); B^* is atmospheric drag coefficient.

All of the parameters listed above are Brouwer mean elements [26] except n_0 . n_0 is a Kozai format [27], which needs transforming into Brouwer format before SGP4 model is used for orbit prediction [28]:

$$\begin{aligned} a_0 &= \left(\frac{k_e}{n_0}\right)^{\frac{2}{3}}, & \delta_0 &= \frac{3k_2}{2a_0^2} \frac{3\cos^2 i_0 - 1}{(1-e_0^2)^{\frac{3}{2}}}, & a'_0 &= a_0 \left(1 - \frac{1}{3}\delta_0 - \delta_0^2 - \frac{134}{81}\delta_0^3\right) \\ \delta_1 &= \frac{3k_2}{2a_0^2} \frac{3\cos^2 i_0 - 1}{(1-e_0^2)^{\frac{3}{2}}}, & n''_0 &= \frac{n_0}{1+\delta_1}, & a''_0 &= \left(\frac{k_e}{n''_0}\right)^{\frac{2}{3}} \end{aligned} \quad (7.9)$$

where $k_e = \sqrt{GM_E}$, where G is a constant of Isaac Newton's gravitation, M_E is the Earth mass; a_0 , δ_0 , a'_0 and δ_1 are derived intermediate parameters; in equation $k_2 = J_2 a_E^2 / 2$, $J_2 = 1.0826261 \times 10^{-3}$ is the potential function term J_2 of the Earth gravity while $a_E = 6371137$ m is the equatorial radius of earth; n''_0 is the mean angular speed of Brouwer format at the epoch time; and a''_0 is the mean semi-major axis of Brouwer format at the epoch time.

For LEO objects, the J_2 term of nonspherical perturbation has no long-term effect upon the semi-major axis and, as its orbit altitude is low, the secular perturbation of the third body attraction and sunlight pressure upon semi-major axis can be neglected. The dominant factor that influences the secular variation of the semi-major axis is atmospheric drag. Though the atmosphere is extremely thin at such altitude, the semi-major axis is still reduced because the space object flies at a rapid speed and for a long time and the atmospheric drag is a dissipative force.

Under the influence of the atmospheric drag perturbation, the secular variation of the semi-major axis can be described in the equation below [28]:

$$a = a_0'' \left[1 - C_1(t - t_0) - D_2(t - t_0)^2 - D_3(t - t_0)^3 - D_4(t - t_0)^4 \right]^2 \quad (7.10)$$

where $t-t_0$ is the prediction interval and C_1 , D_2 , D_3 , and D_4 are related coefficients. See Ref. [28].

For the orbits not synchronous with the sun in LEO objects, the inclination angle is less influenced by the long-term magnitudes of nonspherical perturbation, atmospheric drag perturbation, and the third body attraction perturbation, and the secular variation of the inclination angle is approximate to zero. For sun-synchronous orbits, because that the procession angular speed of RAAN is identical with the angular speed of earth revolution, the relative position of the sun and the orbital plane remains unchanged. Therefore, the solar gravitation can lead to secular variations in orbit inclination angle, which can be expressed in the equation below [29]:

$$\frac{di}{dt} = -\frac{3n_s^2}{16n} \sin i (1 + \cos i_s)^2 \sin(2u_s - 2\Omega) \quad (7.11)$$

where n_s is the mean angular speed of earth revolution; n is the angular speed of orbital motion; i_s is obliquity of the ecliptic; u_s is the mean longitude of the sun over the equator; and Ω is the RAAN of the space object.

7.3.2 Generation and Outlier Detection of Dispersion Data

1. Orbit prediction

The basic method for orbit prediction is as follows: TLE data σ_{k-1} at the time t_{k-1} was predicted with the SGP4 secular term model to obtain the expectant value $\sigma_{k/k-1}$ of secular variation of the orbital element at the t_k time. After that, $a_{k/k-1}$ and $i_{k/k-1}$ were extracted from the expectant value $\sigma_{k/k-1}$. In the meantime, a_k and i_k were acquired based on the cataloging element σ_k at t_k . See Fig. 7.11.

2. Data smoothing and dispersion generation

Because of measurement, orbital determination, and model precision (such as earth, atmosphere, and prediction models), some errors would exist in the cataloged and predicted data to make the catalog and predicted values jitter around the truth value. See Fig. 7.12.

Such jitter could increase the dispersion data so as to cause false alarm under normal circumstances. Thus, a cubic spline curve was used to smooth the data so as to eliminate this small jitter and keep the original characteristic parameter changing in the way it should be. Below is the equation for fitting objective functions. The first term implies fitting offset and the second one denotes fitting smoothness.

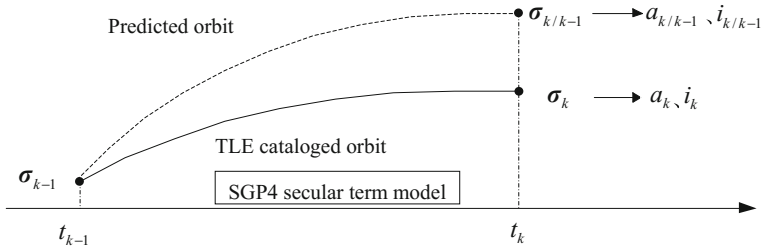


Fig. 7.11 Schematic diagram of orbit prediction based on SGP4 secular term model

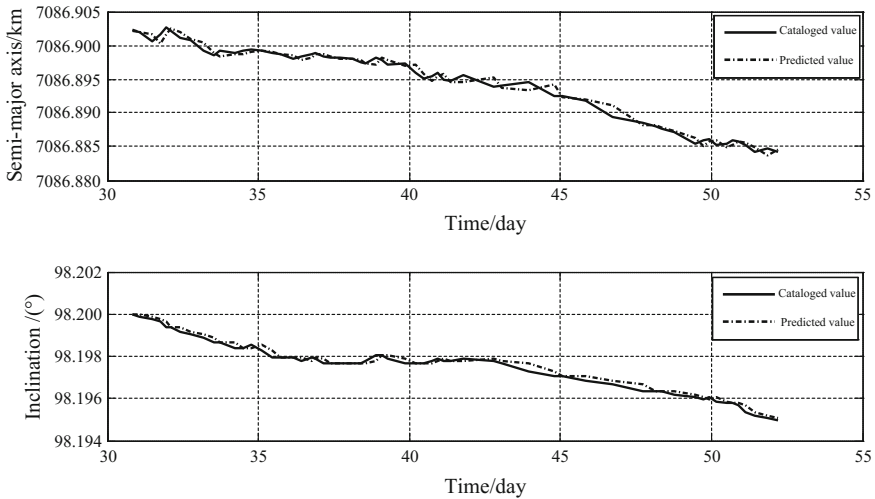


Fig. 7.12 Jitter of characteristic orbit parameters

$$\delta = \omega \sum_{k=1}^n [X(k) - f(t_k)]^2 + (1 - \omega) \int_L [f''(t)]^2 dt \tag{7.12}$$

where X is the data sequence to be smoothed, f is the smoothed function, L is the minimum value in all the epoch time intervals, ω is the weight coefficient with a data range of $[0, 1]$ that decides the data fitting effect.

The error of fitting decreases as ω increases, but the smoothing are less effective. On the contrary, the fitted curve is smoother as ω becomes smaller, but the fitting dispersion increases. Hence, proper ω values could be selected based on different orbit variations to achieve better fitting effect on the promise of smaller fitting dispersions. In general, for drastic transient variations that happened instantaneously, the fitting dispersion should be controlled first, viz., large ω should be

chosen to get a better fitting effect. The ω value was determined with the empirical equation below:

$$\omega = \frac{1}{1 + \bar{t}^3/6}, \quad \bar{t} = \frac{1}{n-1} \sum_{j=1}^{n-1} t_j \tag{7.13}$$

where t_j is the j th time interval and \bar{t} is the mean value of all the epoch time intervals.

Based on the objective function, a cubic spline curve was used to smooth the cataloged and predicted values of the semi-major axis and orbit inclination angle in Matlab to obtain the sequence of dispersion data.

$$\{t_k, \Delta a_k, \Delta i_k\}, \quad k = 1, 2, \dots, n$$

In the above equation,

$$\Delta a_k = a'_k - a'_{k/k-1}, \quad \Delta i_k = i'_k - i'_{k/k-1}, \quad k = 1, 2, \dots, n$$

where a'_k and i'_k are smoothed cataloged values of characteristic orbit elements; $a'_{k/k-1}$ and $i'_{k/k-1}$ are smoothed predicted values of characteristic orbit elements.

The dispersion data contains two parameters—semi-major axis and inclination angle, which can be deemed as multidimensional data anomalies to be detected with Mahalanobis Distance.

The concrete algorithm flow of the above process is specified in Fig. 7.13.

7.3.3 Examples

Terra is an American satellite in sun-synchronous orbit. The five sensors it carries can gather necessary information for observation of the Earth, such as earth atmosphere, lands, oceans, and solar energy balance. This satellite been taken as an

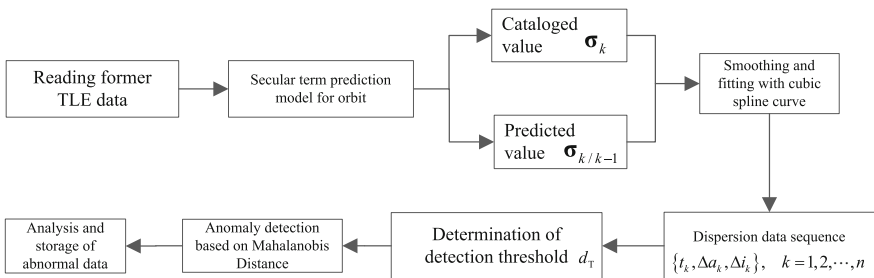


Fig. 7.13 Detection flow of orbital anomalies based on prediction dispersion

Table 7.1 Orbital maneuvers of satellite Terra In 2010

S/N	Date	Time	Maneuver	Orbit Laps	Duration
1	2010-01-22	15:35:00	Altitude maintenance#57	53,711	00:00:06.7
2	2010-02-24	16:12:09	Orbit inclination maneuver #25	54,191	00:05:20.0
3	2010-04-28	13:37:10	Altitude maintenance#58	55,107	00:00:07.7
4	2010-08-11	15:11:00	Altitude maintenance#59	56,637	00:00:05.5
5	2010-10-21	15:28:26	Orbit inclination maneuver#26	57,671	00:05:20.0
6	2010-10-28	15:34:33	Orbit inclination maneuver #27	57,773	00:05:20.0

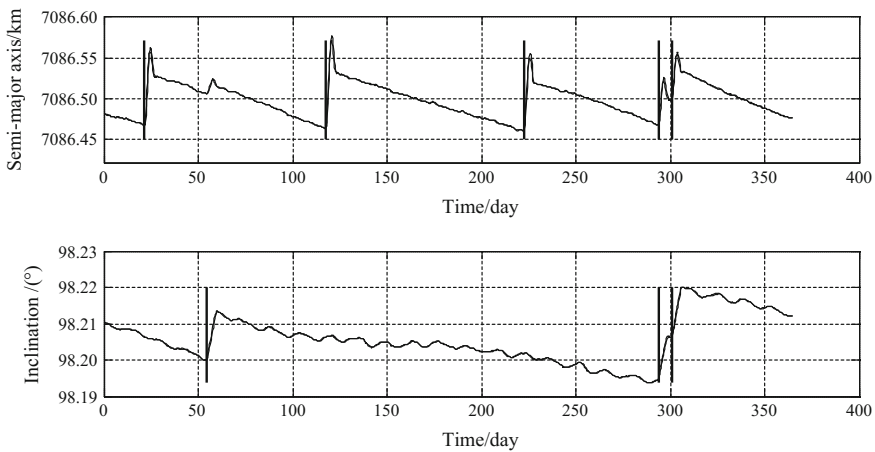


Fig. 7.14 Variations and maneuvers of Terra’s semi-major axis and inclination in 2010

example to show the analysis and detection of orbital anomalies with the method introduced in this section. Table 7.1 shows the historical data of Terra orbital maneuvers made known in 2010 [30].

For the satellite Terra, the TLE data NORAD released in 2010 can be used to analyze the variations of the semi-major axis and inclination angle in 2010, as shown in Fig. 7.14, where the maneuvers in 2010 are marked in vertical lines.

The predicted values of the characteristic orbit elements can be obtained by processing the orbital data of satellite Terra in 2010 with the prediction model referred to in Sect. 7.3.2. The cataloged and predicted values can be processed with the data smoothing method to obtain smoothed values. To make it particular, elaborations have been given below on the processing condition of Terra’s semi-major axis data from the 222nd to the 228th days of 2010 (see Fig. 7.15). The same applies to its orbit inclination angle.

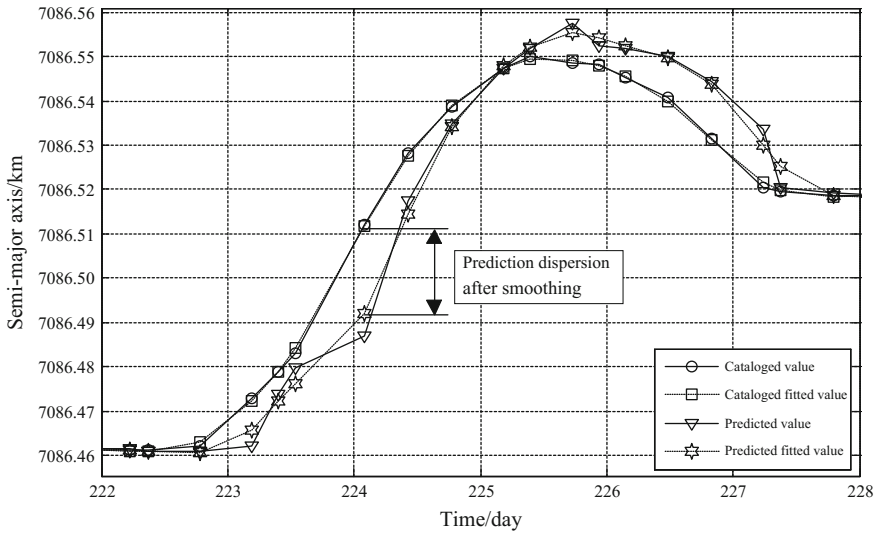


Fig. 7.15 Smoothing of cataloged and predicted values of Terra’s semi-major axis in 2010

The dispersion data and Mahalanobis Distance at each epoch time were obtained after Terra’s semi-major axis and orbit inclination angle had been predicted and smoothed. The number 5 was set as the threshold for anomaly detection of Mahalanobis Distance. See Figs. 7.16 and 7.17 for detection results.

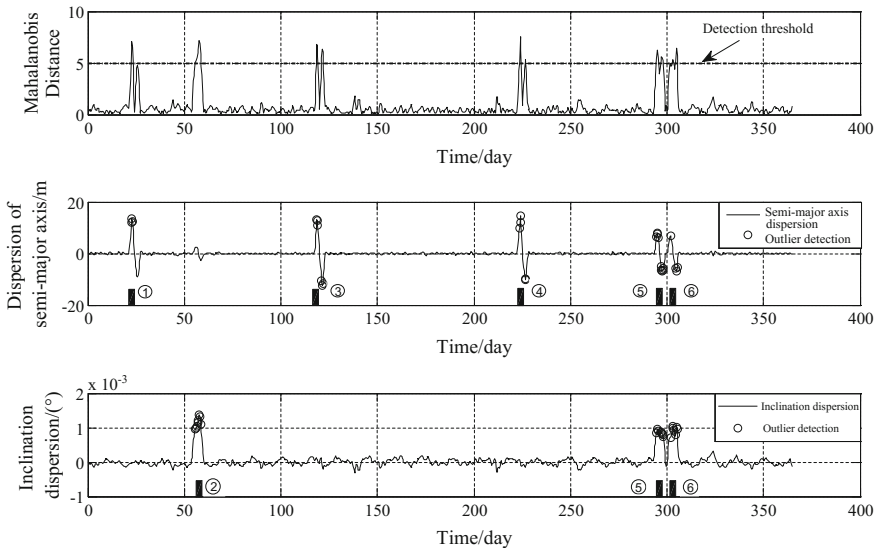


Fig. 7.16 Detection results of Terra maneuvering in 2010

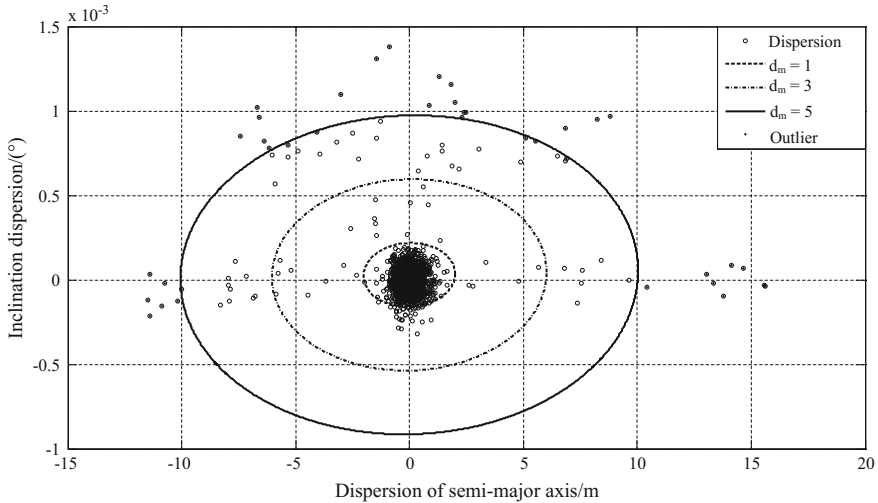


Fig. 7.17 Dispersion distribution of Terra’s semi-major axis and inclination in 2010

As indicated in figures above, the outliers are accumulated near the six time points mentioned in Table 7.1. The reasons are as follows:

- (1) Generally, orbit maintenance is not regulated at a time; instead, it is adjusted and measured continuously within a period of time to reach the desired value.
- (2) TLE orbital determination data is affected by historical data. In the event of orbital anomalies, precise orbital determination is not achieved at a time. On the contrary, it is done with multiple arrays of new orbital data.

From the above contents, we can know that six orbital anomalies happened in 2010. The detected frequency and time of the anomalies are identical with the known orbital maneuver history. We can also know, through analysis, that the first, third, and fourth maneuver are for altitude maintenance and the second, fifth, and sixth maneuver are for orbit inclination maneuver. What’s more, the orbit altitude is enhanced simultaneously at the time when the orbit inclination angle is maneuvered.

7.4 GEO Orbit Anomaly Detection Based on Dispersion of Drift Rate

For the special perturbation and motion principles of GEO objects, the theories concerning mean longitude and inclination angle and the dispersion of actual drift rate have been used in this section to detect GEO orbit anomalies based on the analysis of drifting properties of GEO objects. First of all, the general procedures

were proposed for orbit anomaly detection. Then, mean longitude and inclination angle were selected as characteristic orbit parameters and their variation patterns theoretically analyzed. In the calculation of theoretical and actual drift rates, historical orbital data were smoothed with cubic spline curves. The outliers of dispersion data obtained have been detected and the biasness thereof analyzed. The last step was verification of the method with historical orbital data of two GEO objects.

7.4.1 Characteristic Orbit Parameters and Variation Law

These procedures are also applied to the sections below, especially for orbital anomaly detection of GEO objects based on the dispersion of drift rate. For the convenience of introduction, we have exemplified the Beidou-1B Satellite from China’s Beidou Navigation Satellite System and the failed communication satellite STTW-1 of China. The basic information and orbital parameters of each satellite have been listed in Table 7.2. All the historical TLE data of the two satellites during the time from orbital acquisition to November 2010 have been selected.

1. Selection of Characteristic Orbit Parameters

GEO objects have three features: (1) Its orbital period tallies with the period of earth rotation; it is an earth-synchronous orbit; (2) The orbit is circular with an eccentricity of $e = 0$; (3) The orbit is located on the equatorial plane of the earth to form an inclination angle of $i = 0$. Its RAAN Ω , argument of perigee ω and true anomaly f are indeterminate. Mathematically, they are just singular values. When using classical orbital elements to describe the orbit properties of GEO satellites, there exist singularity problems. To overcome difficulties like small eccentricity and small inclination, these values could be expressed in synchronous orbital element [31–34]. The elements of synchronous orbits have been defined as mean longitude of satellite λ , drift rate of mean longitude D , 2D inclination angle vector i and eccentricity vector e . The relationship between synchronous orbit elements and Kepler orbit elements is displayed as below [31].

$$\begin{aligned}
 \lambda &= s - G(t) & D &= -\frac{3}{2} \frac{a-A}{A} \\
 i_x &= \sin i \sin \Omega & i_y &= -\sin i \cos \Omega \\
 e_x &= e \cos(\omega + \Omega) & e_y &= e \sin(\omega + \Omega)
 \end{aligned}
 \tag{7.14}$$

Table 7.2 Orbital parameters of space objects

SSN No.	Name	Inclination (°)	Perigee (km)	Apogee (km)	TLE Array
26643	Beidou-1B	3.23	35,779	35,794	3586
16526	STTW-1	14.07	35,774	35,812	4037

where s is the sidereal hour angle (SHA) of the object; $G(t)$ is Greenwich hour angle (GHA); the nominal semi-major axis of GEO is $A = 42165.7$ km.

The drift rate of mean longitude has reflected that the dispersion between the orbital period and the earth rotation period, which is dimensionless. In applications, it is generally multiplied by the mean angular speed of earth rotation $\omega_E = 360.985647^\circ/d$ to be converted into the unit of ($^\circ/d$). The vectors of 2D inclination i and eccentricity e have been respectively defined as follows:

$$i = \begin{bmatrix} i_x \\ i_y \end{bmatrix} = \begin{bmatrix} \sin i \sin \Omega \\ -\sin i \cos \Omega \end{bmatrix}, \quad e = \begin{bmatrix} e_x \\ e_y \end{bmatrix} = \begin{bmatrix} e \cos(\Omega + \omega) \\ e \sin(\Omega + \omega) \end{bmatrix} \quad (7.15)$$

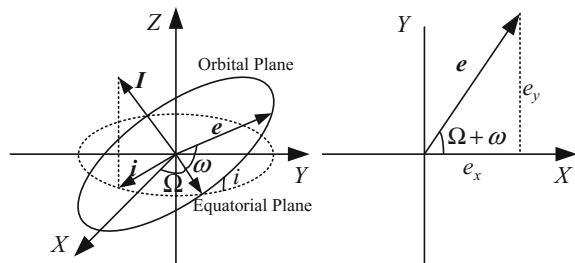
The 2D inclination vector i is the projection of the normal unit vector of a orbital plane on the equatorial plane; eccentricity vector e is a vector in the equatorial plane, which represents the perigee that is pointed from the center of the coordinate system to the orbit. See Fig. 7.18.

Because of perturbations, GEO objects would gradually stray from the nominal position. The maneuver of GEO satellite is mainly for position maintenance, including east-west position maintenance (or so-called longitude position maintenance) and south-north position maintenance (or so-called inclination position maintenance). As the factors to be changed in general for GEO objects, mean longitude λ and inclination angle i were selected as characteristic orbit parameters. Whenever satellite orbits are controlled, the controlled quantity is always calculated according to the secular variation of orbital elements. Thus, mean elements were taken for both the current orbit and target object. The orbital elements given in TLE are mean elements and hence can be directly used to detect orbital anomalies.

2. Position drifting caused by orbital element dispersion

If the orbital elements dissatisfy the requirement on motionlessness, i.e., when small dispersions exist in semi-major axis, eccentricity, and inclination, the position of the object will drift away from where it is. The dispersion of semi-major axis will give rise to long-term drift in longitudinal direction. When dispersion Δa occurs, the drift rate of mean longitude is as follows [31]:

Fig. 7.18 Definition and components of inclination vector and eccentricity vector



$$D_{\lambda} = -\frac{3}{2} \frac{\Delta a}{A} \omega_E \quad (7.16)$$

where the minus sign “-” implies that the object drifts west when $\Delta a > 0$ while the satellite drifts east when $\Delta a < 0$.

If the eccentricity is not zero, it will cause longitudinal swing at an orbital period by the amplitude of $2e$, i.e.,

$$\lambda = \lambda_0 + 2e \sin M \quad (7.17)$$

If the inclination is not zero, it will cause the latitudinal swing at an orbital period by the amplitude of i and the longitudinal swing at a half orbital period by the amplitude of $i^2/4$. Because of this, the track of the satellite position is with a shape of “8” approximately under the influence of the oscillation.

$$\delta = i \sin u, \quad \lambda = \lambda_0 - \frac{i^2}{4} \sin 2u \quad (7.18)$$

The curves of satellite position will become complicated under the influence of the three dispersions in semi-major axis, eccentricity, and inclination.

3. GEO perturbation analysis

As stated above, the position would drifts regularly when the orbital elements were inconsistent with the nominal GEO elements in number. Why the deflection happened will be discussed based on the aspect of perturbation in this subsection. Speaking of GEO, the perturbation terms that need consideration include non-spherical gravity, lunisolar gravitational perturbation, and solar radiation pressure perturbation. Under these perturbative actions, the following parameters change constantly with time: semi-major axis, period, eccentricity, RAAN, and orbit inclination. The GEO gradually deviates from the ideal geostationary orbit with the continuous accumulation of perturbative acceleration.

The drift of GEO parameters is related to three aspects: mean longitude λ , inclination vector i , and eccentricity vector e . Under the first-order approximation, it can be regarded as the followings: (1) The mean longitude λ drifts only under the perturbative action of tesseral harmonic terms of gravitational attraction; (2) The inclination vector i drifts under lunisolar attraction and zonal harmonic terms of gravitational attraction; (3) The drift of eccentricity vector e is mainly driven by the perturbation of solar radiation pressure and slightly by lunisolar attraction. Thus, our main concern lies in the drifting patterns of mean longitude and inclination vector.

(1) Longitude drift

The tesseral harmonic terms of non-spherical gravitation perturbation are caused by asymmetric mass distribution of the earth's interior. GEO satellites with position at nominal precision are subject to additional tangential gravitational acceleration

due to the ellipsoidal shape of the equator. Although the tangential component of acceleration is small in absolute value, yet it is crucial for longitude drift changes. The tangential acceleration is influenced by tesseral harmonic terms C_{22} and S_{22} . Based on the Lagrange's equation, it will make the GEO semi-major axis change to break the conformity between the angular speed of orbital mean motion and the rotational angular speed of the earth, eventually causing the satellite to deviate from the nominal longitude of the satellite.

In GEO synchronous elements, the mean longitude λ and the drift rate D constitutes a drift rate vector (λ, D) . The equation for semi-major axis perturbation can be obtained from the analysis of non-spherical perturbation [34]:

$$\dot{a} = -\frac{2}{3} \frac{A}{\omega_E} \Gamma_\lambda \quad (7.19)$$

where

$$\Gamma_\lambda = -3\omega_E^2 \begin{bmatrix} -6J_{22} \left(\frac{R_E}{A}\right)^2 \sin 2(\lambda - \lambda_{22}) \\ + \frac{3}{2} J_{31} \left(\frac{R_E}{A}\right)^3 \sin(\lambda - \lambda_{31}) \\ -45J_{33} \left(\frac{R_E}{A}\right)^3 \sin 3(\lambda - \lambda_{33}) \end{bmatrix} \quad (7.20)$$

Hence we can get the following equation for the drift rate vector (λ, D) :

$$\begin{cases} \dot{\lambda} = D \\ \dot{D} = \Gamma_\lambda \end{cases} \quad (7.21)$$

From the above we can know that the acceleration of mean longitude perturbation is only associated with the longitude of the satellite. Equation (7.21) is actually a system of nonlinear differential equations.

There are four equilibrium points on the GEO position, that is, an object is firmly fixed on the four positions. The semi-major axis remains unchanged and the acceleration of mean longitude perturbation is zero. Two of them are stable equilibrium points (75.05°E and 104.91°W. These values differ slightly in different works because of different gravitational field models cited. The gravitational field model used in this book is JGM-3). Therefore, the object with the longitude slightly differing from the equilibrium point can still drift back to this point. If the two points are unstable equilibrium points (11.41°W and 162.08°E), satellites will drift from them. If it is situated on other longitudinal positions, GEO object will move with acceleration toward the nearest equilibrium point. If no position maintenance measures are taken, the longitude of GEO object will ultimately waver symmetrically around the stable point with a period of about 800 days.

The expression for the drift rate of mean longitude has been given in Eq. (7.16), where the drift rate of mean longitude is known to depend on the semi-major axis only, while the drift rate of semi-major axis is caused by the tesseral harmonic terms of the non-spherical gravity. Regarding the semi-major axis, long periodic

perturbation is not generated by lunisolar gravitational perturbation. The variation in a cycle is zero, but a short-cycle perturbation equivalent to half a solar day is generated. The non-zero eccentricity and orbit inclination can also cause longitudinal vibrations, but the period of such vibration is smaller or equals to an orbital period (about one day). The release time of TLE elements is generally at the ascending node, it is therefore hard to reflect the periodic longitudinal variation in TLE. So, mean longitude variations are still used for detecting orbital anomalies.

(2) Drift of the inclination vector

The inclination vector is driven to drift both by lunisolar attraction and zonal harmonic terms of the earth's gravity. The latter produces normal perturbation acceleration, which, by changing the RAAN, causes the inclination vector $\mathbf{i} = (i_x, i_y)$ to vary periodically [34]. Such change is reflected by the inclination vector \mathbf{i} revolving with the Z-axis at the speed of $-4.9(^{\circ})/\text{year}$.

The lunisolar attraction is the main perturbative source of GEO inclination drift. The long periodic perturbation of the inclination vector caused by the lunisolar attraction is shown below:

$$\begin{cases} \frac{di_x}{dt} = \frac{3}{8} \frac{n_k^2}{n} \begin{pmatrix} -\sin \Omega_k \sin 2i_k \\ + 2 \sin i_k \cos \Omega_k \sin 2\lambda_k \\ + \sin 2i_k \sin \Omega_k \cos 2\lambda_k \end{pmatrix} \\ \frac{di_y}{dt} = \frac{3}{8} \frac{n_k^2}{n} \begin{pmatrix} \cos \Omega_k \sin 2i_k \\ + 2 \sin i_k \sin \Omega_k \sin 2\lambda_k \\ - \sin 2i_k \cos \Omega_k \cos 2\lambda_k \end{pmatrix} \end{cases} \quad (7.22)$$

where k can be used to indicate the Sun and the Moon.

The perturbation equation above contains constant terms (only i_k and Ω_k) and periodic terms (λ_k). The inclination perturbation has a linear term and a periodic term of time. The period is half of the orbital period of the third body. For lunisolar attraction, there is a half-month periodic term (13.6 days) and a half-year periodic term (182.5 days).

The general influence of all the perturbations upon the inclination vector can be expressed with a mean approximate, that is, the inclination vector \mathbf{i} revolving around a pole $(0, -7.4^{\circ})$ in a negative direction with a period of 54 years.

The orbital plane (moon's path) in which the moon moves around the earth processes against the ecliptic pole for a period of 18.6 years. If we set Ω_m as a celestial longitude of the moon's path on the ascending node of the ecliptic, then we can get the equation below:

$$\Omega_m = 259.183^{\circ} - 0.05295^{\circ} \cdot t \quad (7.23)$$

where t is a Julian day calculated from January 1st, 1900 at 12:00 a.m.

Then, the long periodic perturbation of inclination is expressed as follows [29]:

$$\begin{aligned} \frac{di_x}{dt} &= 0.859 + 0.107 \cos \Omega_m - 0.0025 \cos 2\Omega_m \\ \frac{di_y}{dt} &= 0.134 \sin \Omega_m - 0.003 \sin 2\Omega_m \end{aligned} \tag{7.24}$$

where the unit of the drift rate is (°)/year.

Based on the approximate expression of orbit inclination $i^2 = i_x^2 + i_y^2$, the drift rate of the orbit inclination can be calculated as follows:

$$D_i = \frac{di}{dt} = \frac{1}{i} \left(i_x \frac{di_x}{dt} + i_y \frac{di_y}{dt} \right) \tag{7.25}$$

As we have known the values of semi-major axis a , inclination i , and RAAN Ω , the theoretical drift rates of mean longitude and inclination can be calculated according to Eqs. (7.16) and (7.25). See Figs. 7.19 and 7.20 for the theoretical drift rates of mean longitudes and inclination angles of Beidou-1B and STTW-1 Satellites.

7.4.2 Generation of Dispersion Data of Drift Rate

1. Preprocessing of orbital data

The dispersion data of drift rate is generated by calculating the theoretical and actual drift rates of characteristic orbit parameters (mean longitude and inclination). The mean element data of semi-major axis a are needed to calculate the theoretical

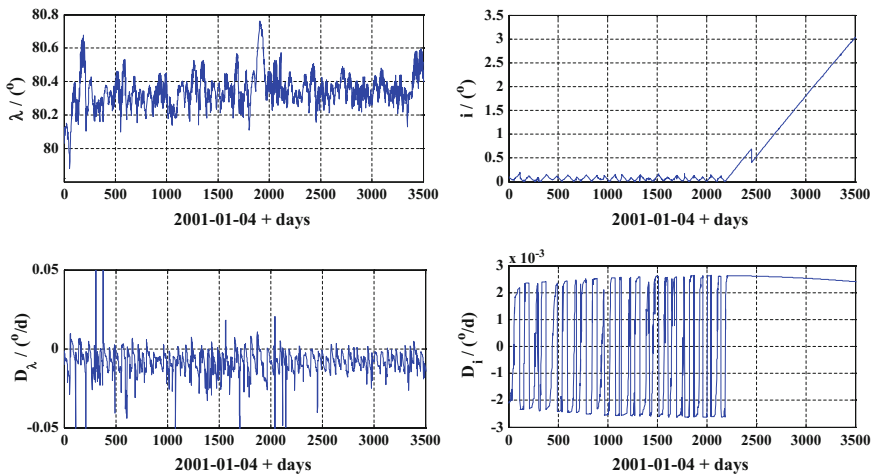


Fig. 7.19 Theoretical drift rates of mean longitude and inclination of Beidou-1B satellite

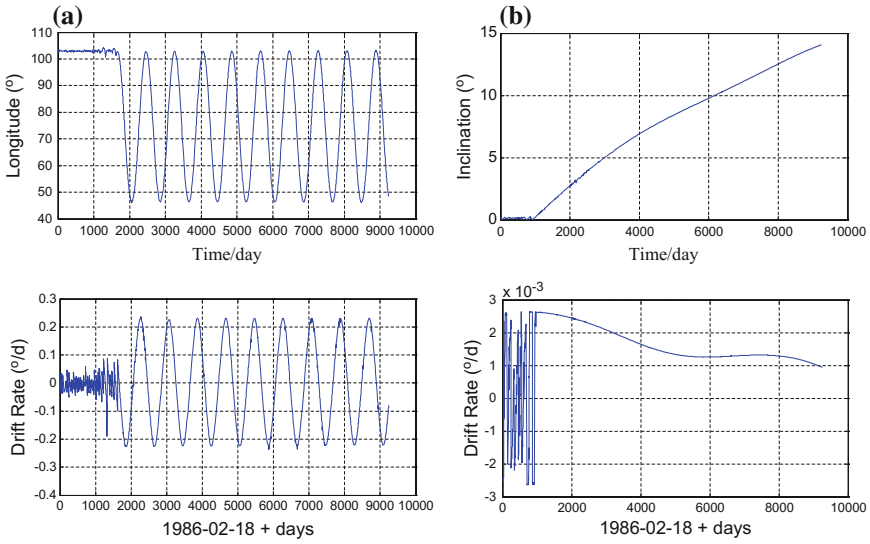


Fig. 7.20 Theoretical drift rates of mean longitude and inclination of STTW-1 satellite

drift rate of mean longitude, and the mean elements of inclination i and RAAN Ω are needed to calculate the theoretical drift rate of the inclination. The data of mean longitude λ and inclination i are needed to calculate the actual drift rates of mean longitude and inclination. All these data are selected from historical TLE data of space objects. Thus, errors can hardly be avoided, as is the case for theoretical and actual drift rates. In Fig. 7.21, noise exists in the semi-major axis data of Beidou-1B Satellite and there are notable outliers at some points. If this data is directly used to calculate the theoretical drift rate of men longitude, large errors will be produced.

To solve this problem, the historical orbital data should be processed to inhibit the noise. In this section, the orbital elements have been processed by cubic spline smoothing function CSAPS from Matlab R2010b in the calculation of theoretical and actual drift rates. The cubic spline smoothing is a data processing method that uses spline functions to fit the data with noise. The spline function need to be properly constructed to minimize the fitted index function, i.e.,

$$F = p \sum_{i=1}^n [y_i - f(t_i)]^2 + (1 - p) \int [f''(t)]^2 dt \tag{7.26}$$

The spline function f is constructed based on a specified smoothing parameter p . In the equation above, the first term means the fitting deviation and the second one denotes fitting smoothness. The smoothing parameter p represents the extent to which the data is smoothed by the spline function f . Its value range is $[0, 1]$. Its smoothness rests on how far the value is away from zero. When $p = 0$, f is the Least Squares linear fitting of the data; the closer the value is to zero the smoother the data

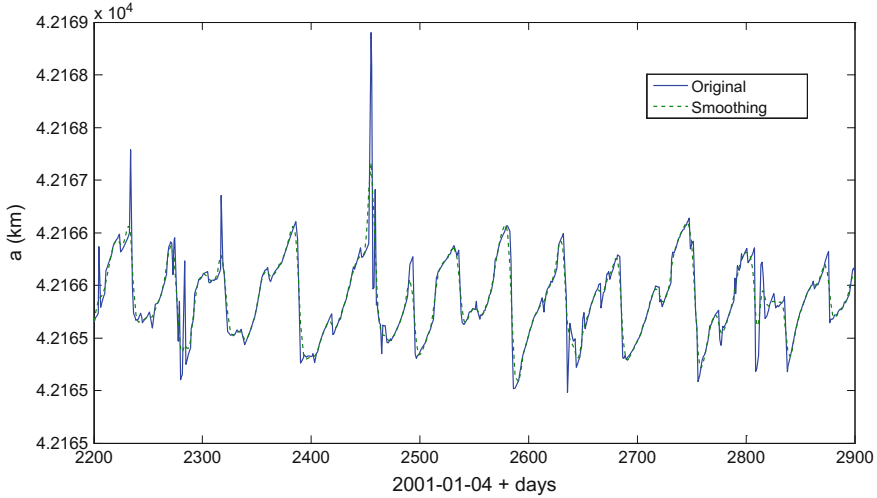


Fig. 7.21 Initial and smoothed data of semi-major axis of Beidou-1B satellite

becomes. When $p = 1$, f is a cubic spline interpolation of the data. As shown in Fig. 7.21, the dotted line stands for the smoothed data of semi-major axis, by which noises have been removed from the original data and the outliers have been remarkably eliminated.

2. Calculation of theoretical and actual drift rates and generation of dispersion

New orbital data have been obtained by using cubic splines to smooth the historical data of semi-major axis a , inclination i , RAAN Ω , and substellar point longitude. These new orbital data can be applied in two aspects. On one side, the three parameters namely a , i , and Ω can be processed with Eqs. (7.16) and (7.25) to calculate the theoretical drift rates of mean longitude and inclination. Figure 7.22 shows the curve of the theoretical drift rate of mean longitude calculated with original and smoothed semi-major axis data of Beidou-1B Satellite. As is indicated in the figure, the results calculated with smoothed data are more stable. On the other side, after the spline functions of substellar point longitude λ and inclination i are calculated, their change rates at each time can be derived and these change rates can be used as approximate values of actual drift rates.

Different smoothing parameters (p) have been selected to calculate theoretical and actual drift rates, respectively. The calculation of theoretical drift rates needs smoothed data whereas the calculation of actual drift rates needs derivatives of functions. Smoothing parameters $p_l = 0.1$ and $p_e = 0.005$ have been taken to calculate theoretical and actual drift rates in respective.

After the aforesaid drift rates have been obtained, subtracting one from the other can get the dispersion at each time and the dispersion data sequence $\{t^{(k)}, \Delta D_\lambda^{(k)}, \Delta D_i^{(k)}\}$, $k = 1, 2, \dots, n$.

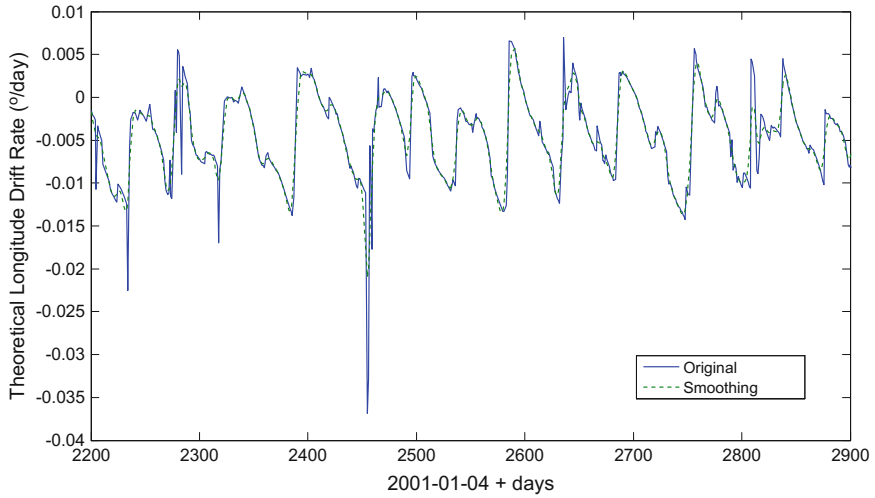


Fig. 7.22 Theoretical drift rate of mean longitude of Beidou-1B satellite calculated based on initial and smoothed data

Figures 7.23 and 7.24 are 2D distributions for the dispersion data of mean longitude and inclination of satellites Beidou-1B and STTW-1. The dispersion can be used as detectable quantity of orbital anomalies.

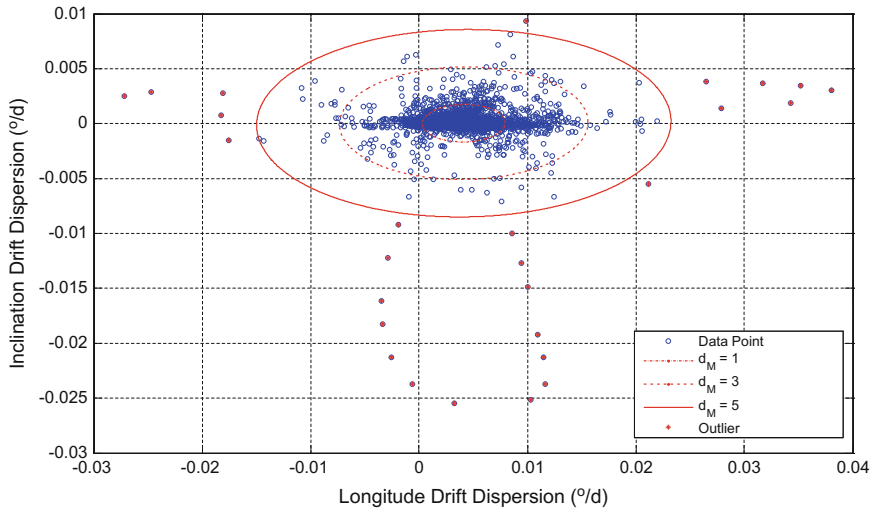


Fig. 7.23 Detection of outliers in drift rate dispersion based on Mahalanobis distance (Beidou-1B)

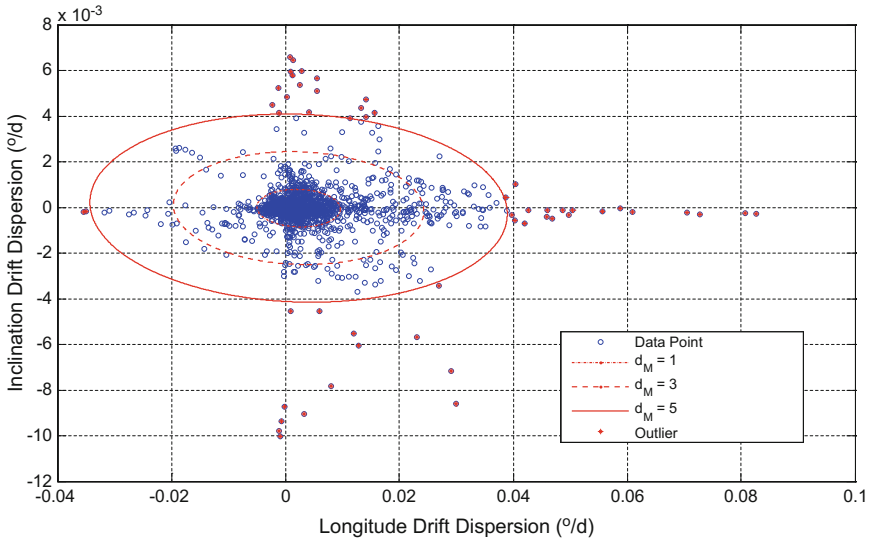


Fig. 7.24 Detection of outliers in drift rate dispersion based on Mahalanobis distance (STTW-1)

7.4.3 Detection of Outliers in Dispersion Data

The detection of orbital anomalies is based on the determination of the calculated actual and expectant values of characteristic parameters. If the dispersion is incredibly large, it is supposed to be abnormal. The dispersion data obtained in the section above can be used to detect orbital anomalies.

After the dispersion data is obtained, the detection of orbital anomalies is actually the detection of outliers in dispersion data. Two characteristic parameters needs to be both considered for the method of detecting orbital anomalies based on the drift rate dispersions of mean longitude and inclination, and their dispersion data must be considered as well in the detection of outliers. Mahalanobis Distance should be used to detect the outliers. In the 2D space, the curves of equal Mahalanobis Distances ($d_M = 1, 3, 5$) are a series of ellipses as drawn in Figs. 7.23 and 7.24. If the threshold value is $d_T = 5$, the data points outside the ellipses of equal Mahalanobis Distances would be regarded as outlier points, which have been highlighted with asterisks in the diagrams below.

According to the detection results of orbital anomalies of Beidou-1B Satellite in Fig. 7.23, if orbital maneuvers were frequented when the satellite remains controllable within the time of analysis, the dispersion of drift rate would be too large to detect the outliers.

Regarding uncontrolled and normal orbits, the distribution of drift rate dispersions should be unbiased, and the dispersion in the drift rate of mean longitude is irrelevant to that of inclination. It is reflected on the dispersion distribution, i.e., the 2D scattering centers of dispersions in drift rates of mean longitude and inclination

are near the origin of the coordinate system. The principal axis of the distribution ellipsoid is parallel to the coordinate axis. The uncontrolled phase of satellite STTW-1 has been selected. Figure 7.25 shows the distribution of drift rate dispersions at the uncontrolled phase within 6000 days from April 19th, 1994. As seen in the figure, the eccentric center of equal Mahalanobis Distance for dispersion distribution is near the origin of the coordinate system and the principal axis is parallel to the coordinate axis.

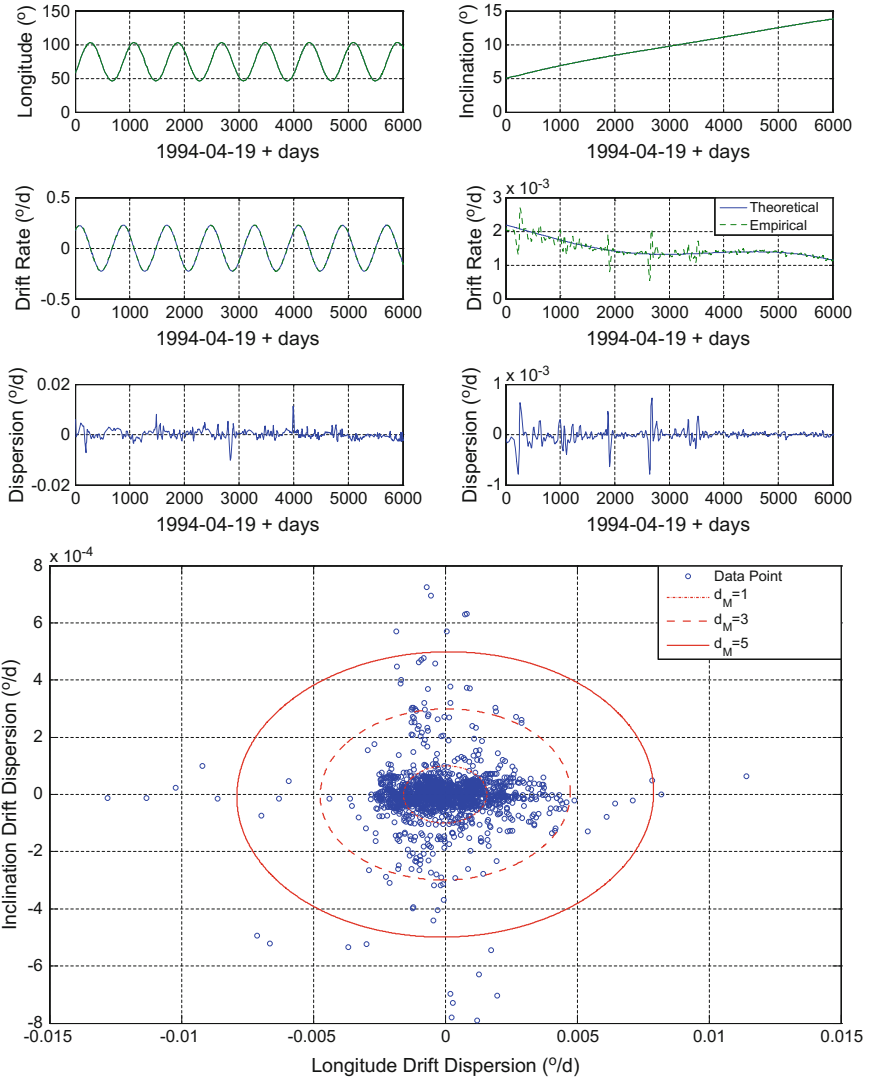


Fig. 7.25 Distribution of drift rate dispersions at uncontrolled sections of STTW-1 satellite

Regarding the satellites controlled within the time of analysis, the dispersions of drift rate are deflected significantly. As shown in Fig. 7.23, the drift rate dispersion of Beidou-1B Satellite is distributed obliquely in the direction of drift rate dispersion of mean longitude. The abscissa for eccentric center of equal Mahalanobis Distance is 0.004 ($^{\circ}$)/day or so. As the drift rate dispersions of uncontrolled orbits are distributed in an unbiased manner, the bias can be used as a foundation for the detection of orbital maneuvers.

As analyzed above, when the outliers are detected according to Mahalanobis Distance, it is determined not with the Mahalanobis Distance from the sample to the distribution center but with the Mahalanobis Distance from the sample to the origin. Figure 7.26 shows the detection results of maneuvers of Beidou-1B Satellite based on the Mahalanobis Distance from the sample to the origin. As we can see from the maps below, most of longitude and inclination maintenance maneuvers can be detected, which is advantageous over the detection results determined with the Mahalanobis Distance from the sample to the eccentric center.

7.4.4 Examples

In conclusion, the detection method for orbital anomalies of GEO objects based on drift rate dispersions is indicated in Fig. 7.28. The detection results for two GEO satellites listed in Table 7.2 are shown in Fig. 7.26. The longitude of Beidou-1B Satellite has always been controllable and its position has been fixed near 80.3°E . The orbit inclination started to drift from the 2190th day. It was maneuvered to stay still on the 2452nd day but drifted again on the 2454th day. As seen in Fig. 7.26, most of the longitude and inclination maintenance maneuvers have been detected. Since the satellite positioning longitude is 80.3°E , the satellite tends to drift west toward 75°E . The dispersion of drift rate is maneuvered to be positive, which coincides with the results in the figure.

The orbital inclination of STTW-1 Satellite started to lose control from about 900th days after the satellite was launched into orbit. It has drifted to 14.067° up to now. The orbital semi-major axis was out of control from about the 1623rd day (June 1990) and oscillated at a period of 800 days (2.19 years) by the amplitude of 17.5 km. The periodic oscillation led to vibration in orbital period, thus forcing the substellar point longitude to swing between 46.5°E and 103°E at the same period. The center of oscillation is a stable equilibrium point at 75°E . As shown in Fig. 7.27, the orbital maneuvers detected were mainly conducted within 2000 days after satellite was launched into orbit, suggesting that in this period, the satellite was controllable and the altitude and inclination has been continuously maneuvered. From the 2000th day to now, no orbital maneuvers have been detected, which indicates that the satellite has been out of control to drift under natural disturbing forces. The maneuver of orbit inclination was detected near the 2000th day and several abnormal changes were found in 400 days from the 2000th day to the 2400th day.

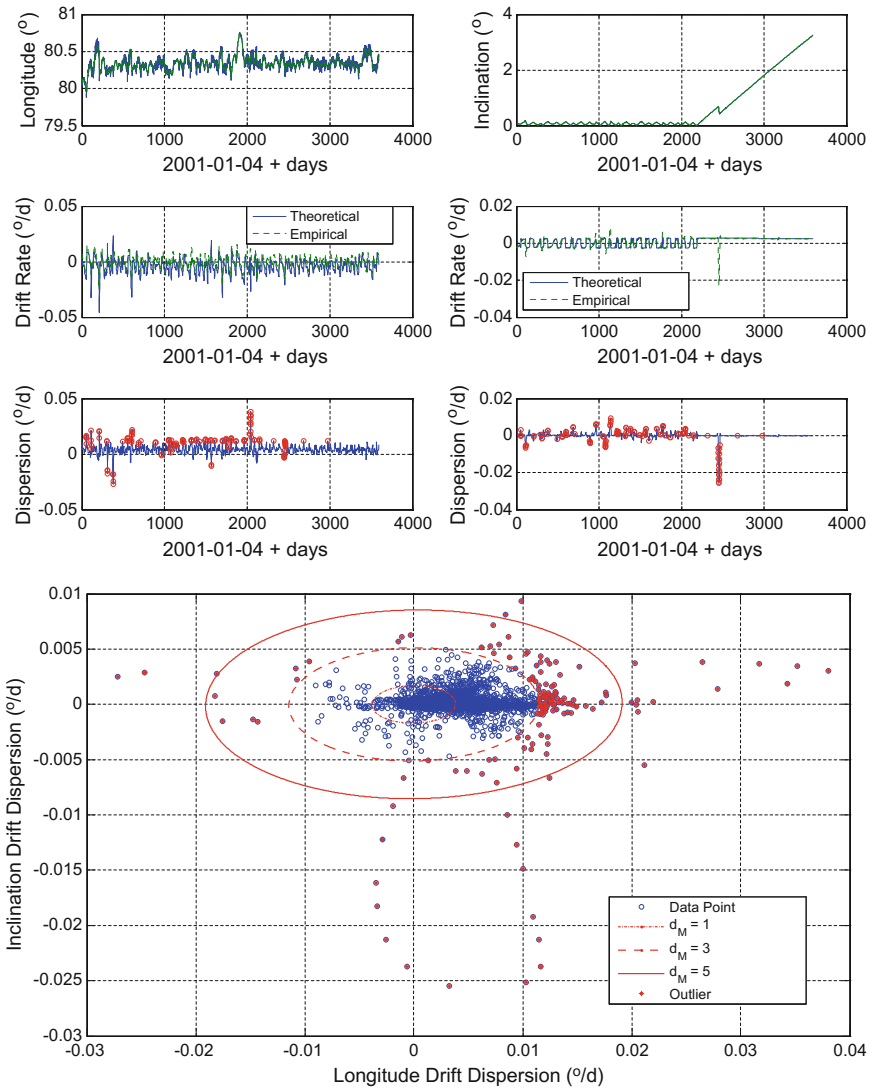


Fig. 7.26 Detection results of orbital maneuvers of Beidou-1B satellite

7.5 Assessment of Detection Method of Orbital Anomaly

Three aspects have been assessed for the detection method of orbit anomalies: demand analysis for orbital precision, parameter sensitivity analysis, and misjudgment probability analysis.

Whether the detection method of orbital anomalies is reliable depends on the prediction and estimation precision of actual and expectant values of the dispersion

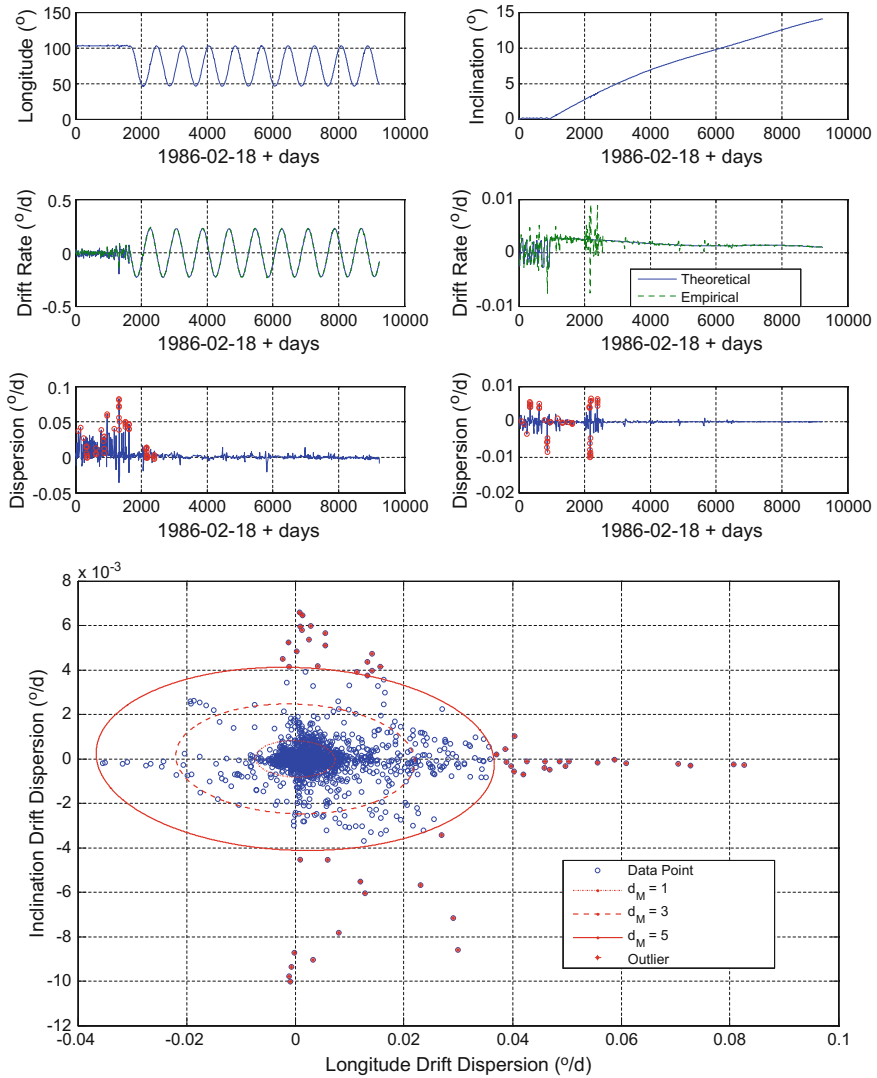


Fig. 7.27 Detection results of orbital maneuvers of STTW-1 satellite

data. The precision of orbital prediction and orbital determination of TLE data has been analyzed in this chapter.

Some configuration parameters are needed to detect orbital anomalies, such as the window size of MWCF, polynomial order, detection threshold, and so on. Any changes in these parameters can affect the detection method.

In fact, to assess the detection method of orbital anomalies is to evaluate its advantages and disadvantages. It discriminates “normal” dispersion data from



Fig. 7.28 Flow chart of GEO orbit anomaly detection based on dispersion of drift rate

“abnormal” ones. As a means of discrimination, two types of misjudgment are always unavoidable: false alarm and missing alarm. It is vital to calculate and estimate the misjudgment probability. Both the false alarm and the missing alarm

need to be controlled, but they are contradictory to each other. When the false alarm probability is definite, the missing alarm probability is lessened as the Signal to Noise Ratio (SNR) increases. When the SNR is certain, the smaller the false alarm probability is, the larger the missing alarm probability becomes, vice versa. In practice, the false alarm probability is always preset while the missing alarm probability is lowered by pushing up the SNR.

In this section, the MWCF is used as an example to analyze the precision of the detection method in orbital determination and prediction, more specifically, to analyze the parameter sensitivity and misjudgment probability of the method under certain orbital precision.

7.5.1 Discriminant Analysis and Misjudgment Probability

The detection of space events virtually concerns the discriminant analysis of dispersion data, namely “normal” from “abnormal” and “maneuvered” from “not maneuvered.” Discriminant analysis is one of widely applied methods in multivariate analysis.

It can be expressed in mathematical linguistics: First, set a population φ that contains m daughters $\Pi_1, \Pi_2, \dots, \Pi_m$. The proportion of each daughter in the population is indicated by $\pi_1, \pi_2, \dots, \pi_m$, $\sum_{i=1}^m \pi_i = 1$. Second, assume that each daughter has its own probability distribution (discrete) or density function (continuous) $f_i(\mathbf{x})$, $i = 1, 2, \dots, m$, wherein \mathbf{x} is a p -dimensional variable vector. Third, judge that the sample \mathbf{x} extracted from the population φ belongs to which daughter. The followings are specific procedures: divide the p -dimensional Euclidean space R^p into m mutually disconnected sub-domains R_1, R_2, \dots, R_m , i.e., $R_i \cap R_j = \emptyset$ ($i \neq j$), $\cup_{i=1}^m R_i = R^p$. For the preset $\mathbf{x} \in R^p$, when $\mathbf{x} \in R_i$, determine that \mathbf{x} is extracted from the daughter $\Pi_i, i = 1, 2, \dots, m$. The division of R^p is called a discrimination rule.

Apparently, to solve a discriminant problem needs to determine if the discrimination rule is “good.” Here, the concept “good” can be either based on a theoretic criterion or an intuitive and feasible idea. Still, misjudgment is unavoidable, no matter what kind of discrimination rule is used. For example, \mathbf{x} is extracted from the daughter Π_i , but mistakenly judged to be extracted from Π_j . The misjudgment probability is expressed below:

$$P_{\text{err}}(j|i) = \int_{R_j} f_i(\mathbf{x}) d\mathbf{x}, \quad i \neq j \quad (7.27)$$

The integral sign in the equation above represents multiple integral (for continuous variable) or summation (for discrete variable). Hence, the total average misjudgment probability of a discrimination rule is as follows:

$$P_{\text{err}} = \sum_{i=1}^m \pi_i \sum_{i \neq j} P_{\text{err}}(j|i) = \sum_{i=1}^n \Pi : \sum_{i \neq j} \int_{R_j} f_i(\mathbf{x}) d\mathbf{x} \quad (7.28)$$

The misjudgment probability should be as small as possible, no matter what kind of discrimination rule is used. In practice, misjudgment probabilities are all estimated because $f_i(\mathbf{x})$ is unknown or partly unknown.

It is important to calculate and estimate misjudgment probabilities in discriminant analysis. Whether a misjudgment probability is large or not is determined by two reasons. For one thing, it is subject to those daughters to be discriminated. If these daughters are difficult to be discriminated from each other, the misjudgment probability may still be high, even though the best discrimination rule is used. For another, it is subject to what kind of method is used. Different discrimination methods applied will lead to different misjudgment probabilities. So, a satisfied discrimination method (with the lowest misjudgment probability) can be chosen by calculating and estimating the misjudgment probability. Besides, if the misjudgment probability cannot be remarkably reduced whatever methods are used, the discrimination itself can be considered “not good” and cannot be used to deduce any significant conclusions.

There are four probabilities in detection of orbital anomalies: The detection probability is P_d when there exist orbital anomalies and they are detected; the missing alarm probability is P_m when there exist orbit anomalies but no one is detected; the false alarm probability is P_{fa} when there does not exist orbit anomaly but orbital anomalies are detected; the normal probability is P_n when there exist no orbital anomalies and also no anomaly is detected. The four probabilities are mutually dependent. The relationship between the detection probability and the missing alarm probability is $P_d + P_m = 1$ and between the normal probability and the false alarm probability is $P_{fa} + P_n = 1$.

In the four cases above, false alarm and missing alarm are both misjudgments, which need to be controlled. P_m and P_{fa} are contradictory to each other. The relationship between P_m , P_{fa} , and α (SNR) is this: When P_{fa} is definite, the larger α is the smaller P_m is; when α is certain, the smaller P_{fa} is the larger P_m is, vice versa. In practice, P_{fa} is often specified and α is increased to reduce P_m .

7.5.2 Probability of False Alarm

1. The 2D dispersion vector for false alarm probability is theoretically calculated and defined as below:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (7.29)$$

If no orbital anomalies exist, the dispersion vector \mathbf{x} will be subject to 2D Gaussian distribution with a mean value $\boldsymbol{\mu}$ and a covariance matrix $\boldsymbol{\Sigma}$, namely $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and the probability density function of \mathbf{x} will be the followings:

$$f(\mathbf{x}) = \frac{1}{2\pi|\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right] \quad (7.30)$$

If we assume that the covariance matrix is a diagonal matrix $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \sigma_2^2)$, Eq. (7.30) then can be written in component form, i.e.,

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\frac{1}{2}\left(\frac{(x_1 - \mu_1)^2}{\sigma_1^2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2}\right)\right] \quad (7.31)$$

So, the probability for 2D random vector \mathbf{x} to fall within the elliptic domain of equal Mahalanobis Distance

$$E: \frac{(x_1 - \mu_1)^2}{\sigma_1^2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \leq d^2$$

can be expressed in the equation below:

$$P(d) = \iint_{\frac{(x_1 - \mu_1)^2}{\sigma_1^2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \leq d^2} \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\frac{1}{2}\left(\frac{(x_1 - \mu_1)^2}{\sigma_1^2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2}\right)\right] dx_1 dx_2 \quad (7.32)$$

After the transformation of coordinate systems,

$$x_1 = \sigma_1 r \cos \theta + \mu_1, \quad x_2 = \sigma_2 r \sin \theta + \mu_2 \quad (7.33)$$

one has

$$P(d) = \frac{1}{2\pi} \int_0^d \int_0^{2\pi} r e^{-r^2/2} dr d\theta = \int_0^d r e^{-r^2/2} dr = 1 - e^{-d^2/2} \quad (7.34)$$

The probability for 2D random vector \mathbf{x} to fall outside the elliptic domain of equal Mahalanobis Distance can be indicated in the equation below:

$$P_{\text{fa}} = 1 - P(d) = e^{-d^2/2} \quad (7.35)$$

Thus, the normal probability and false alarm probability are generally expressed as follows:

$$P_n = 1 - e^{-d^2/2}, \quad P_{fa} = 1 - P_n = e^{-d^2/2} \tag{7.36}$$

Figure 7.29 shows the curve of the normal probability P_n and false alarm probability P_{fa} Changing with the Mahalanobis Distance threshold.

2. Actual calculation of false alarm probability

In Eq. (7.30) above, $(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})$ is the square of Mahalanobis Distance from \mathbf{x} to population. In actual detection, the mean value $\boldsymbol{\mu}$ of dispersion vector and the covariance matrix $\boldsymbol{\Sigma}$ are unknown and can only be approximated by the sample mean of unbiased estimate and the sample covariance matrix, namely

$$\bar{\mathbf{x}} = \sum_{i=1}^n \mathbf{x}_i, \quad \mathbf{C} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \tag{7.37}$$

Therefore, the square of Mahalanobis Distance is expressed below:

$$d_M^2 = (\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{C}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) \tag{7.38}$$

The sample \mathbf{x}_i complies with 2D Gaussian distribution $\mathbf{x}_i \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. In Eq. (7.38), $\mathbf{x}_i - \bar{\mathbf{x}}$ can be written as follows:

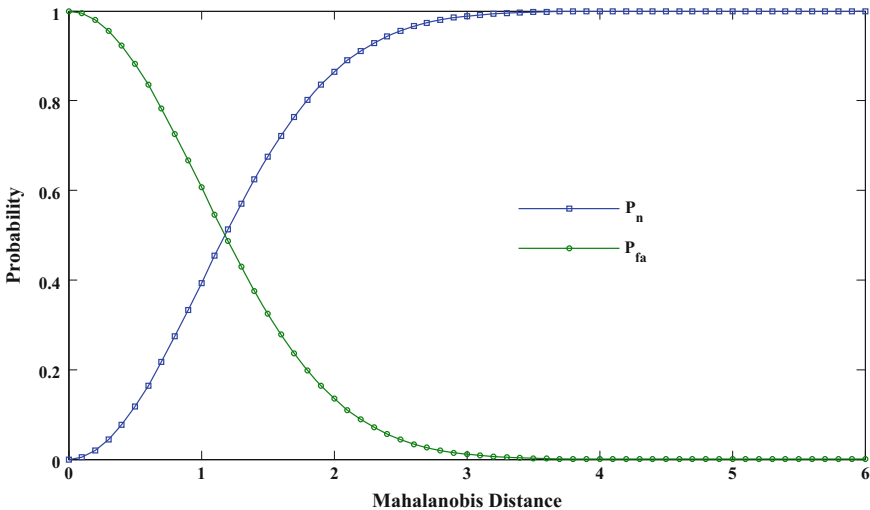


Fig. 7.29 Curve of P_n and P_{fa} changing with Mahalanobis distance threshold

$$\mathbf{x}_i - \bar{\mathbf{x}} = \mathbf{x}_i - \frac{1}{n} \sum_{j=1}^n \mathbf{x}_j = \frac{n-1}{n} \mathbf{x}_i - \frac{1}{n} \sum_{\substack{j=1 \\ j \neq i}}^n \mathbf{x}_j = \frac{n-1}{n} (\mathbf{x}_i - \bar{\mathbf{x}}_{n-1}) \quad (7.39)$$

In the equation above, $\bar{\mathbf{x}}_{n-1}$ is the mean vector of $n-1$ samples, except \mathbf{x}_i that is irrelevant to $\bar{\mathbf{x}}_{n-1}$. $\bar{\mathbf{x}}_{n-1}$ complies with 2D Gaussian distribution $\bar{\mathbf{x}}_{n-1} \sim N_2(\boldsymbol{\mu}, \frac{1}{n-1} \boldsymbol{\Sigma})$, so the random vector $\mathbf{x}_i - \bar{\mathbf{x}}$ complies with 2D Gaussian distribution $\mathbf{x}_i - \bar{\mathbf{x}} = \frac{n-1}{n} (\mathbf{x}_i - \bar{\mathbf{x}}_{n-1}) \sim N_2(\mathbf{0}, \frac{n-1}{n} \boldsymbol{\Sigma})$.

So, the sample covariance matrix $(n-1)\mathbf{C}$ complies with Wishart distribution in which the Degree of Freedom (DOF) is n , namely

$$(n-1)\mathbf{C} = \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \sim W_2\left(\frac{n-1}{n} \boldsymbol{\Sigma}, n\right) \quad (7.40)$$

Hence the following equation is obtained:

$$\frac{n}{n-1} (\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{C}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) \sim T^2(2, n) \quad (7.41)$$

where T^2 was put forward by Hotelling. In multivariate statistics, Wishart distribution is used as the χ^2 distribution in univariate statistics while the function of the latter as the t distribution in univariate statistics.

As indicated in Eq. (7.38), the following formula should be obtained when no orbital anomalies are detected:

$$\frac{n}{n-1} d_M^2 \sim T^2(2, n) \quad (7.42)$$

This shows that, under normal conditions, the square of Mahalanobis Distance $nd_M^2/(n-1)$ complies with distribution $T^2(2, n)$. From the relationship between T^2 distribution and F distribution,

$$\frac{n-m+1}{mn} T^2(m, n) \sim F(m, n-m+1) \quad (7.43)$$

it can be obtained that

$$\frac{1}{2} d_M^2 \sim F(2, n-1) \quad (7.44)$$

Figures 7.30 and 7.31 are curves of probability density function and probability distribution function of d_M obtained when different n values are taken. Theoretical calculation results have been given as well. In the figures, the actual calculation

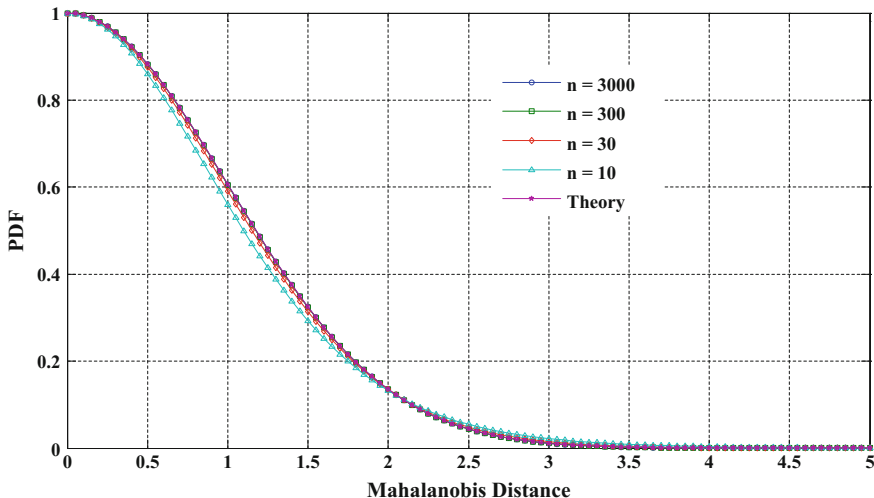


Fig. 7.30 d_M Probability density function in terms of different n values

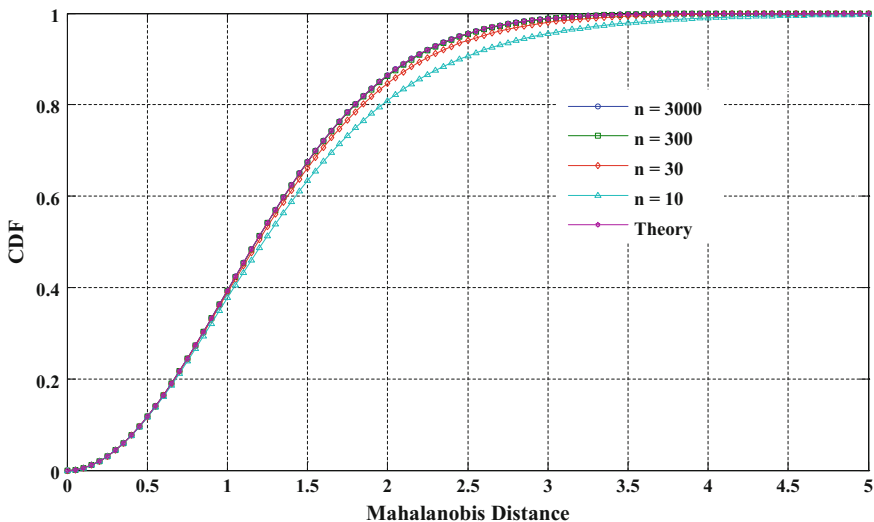


Fig. 7.31 d_M Probability distribution function in terms of different n values

results are closer to the theoretical ones as n increases, and they become the same when $n \rightarrow \infty$.

As n increases, the normal probability tends to increase too while the false alarm probability tends to decrease. The limit value of $n \rightarrow \infty$ is the theoretical calculation value. When $n = 300$, under the condition that the dispersion vector complies with 2D Gaussian distribution, the probabilities that the Mahalanobis Distance from

the sample to its distribution center (weighted by sample covariance matrix) is larger than 1, 3, and 5 are 60.70, 1.187, and 0.0006%, respectively. The results demonstrate that, if the detection threshold of Mahalanobis Distance is set as 3, the probability that the Mahalanobis Distance exceeds the threshold is 1.187% when no anomalies are detected, that is, the false alarm probability is 1.187%. We can thus know the relationships between sampling number, detection threshold, and false alarm probability.

7.5.3 Probability of Missing Alarm

The false alarm probability P_{fa} is always given in practical engineering and the SNR α is enhanced to reduce the missing alarm probability P_m . In the detection of orbital anomalies, noise is the errors of characteristic orbit parameters. If it orbital anomalies really exist, signal is an abnormal orbit parameter. The SNR is the standard deviation ratio of abnormal and normal orbital parameters.

The 2D dispersion vector corresponding to orbital anomalies is assumed to meet 2D Gaussian distribution as well, its mean value is μ_A and the covariance matrix is Σ_A . If a data point belongs to orbital anomaly, it then complies with the distribution $\mathbf{x} \sim N_A(\mu_A, \Sigma_A)$ and the probability density function of 2D dispersion vector \mathbf{x} is as follows:

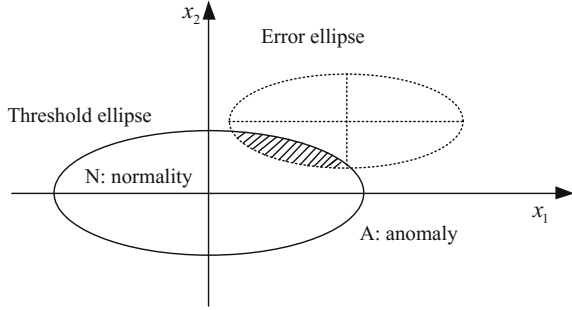
$$f(\mathbf{x}|A) = \frac{1}{2\pi|\Sigma_A|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \mu_A)^T \Sigma_A^{-1}(\mathbf{x} - \mu_A)\right] \quad (7.45)$$

Similarly, if the covariance matrix is set as diagonal matrix $\Sigma_A = \text{diag}(\sigma_{A1}^2, \sigma_{A2}^2)$, Eq. (7.45) can be written as follows:

$$f(x_1, x_2|A) = \frac{1}{2\pi\sigma_{A1}\sigma_{A2}} \exp\left[-\frac{1}{2}\left(\frac{(x_1 - \mu_{A1})^2}{\sigma_{A1}^2} + \frac{(x_2 - \mu_{A2})^2}{\sigma_{A2}^2}\right)\right] \quad (7.46)$$

Figure 7.32 is the schematic diagram of false alarm probability and missing alarm probability. On the 2D plane, the ellipse of Mahalanobis Distance threshold marked in full line divides the 2D plane into two parts: the inner ellipse is zone N, which stands for normality. If the dispersion vector falls within this zone, it is then determined as orbital normality; the outer ellipse is zone A, which denotes anomaly. If the dispersion vector falls within this zone, it is then determined as orbital anomaly. False alarm probability refers to the probability for the dispersion vector to fall within zone A when the orbit is in a normal state. Missing alarm probability means the probability for the dispersion vector to fall within zone N when the orbit is in an abnormal state (as shown in the dispersion vector distribution of abnormal orbits marked in the dashed ellipse), which can be schematically indicated in the dash area.

Fig. 7.32 Probabilities of false and missing alarm



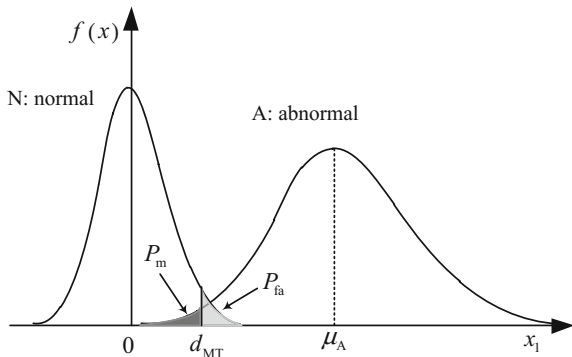
It should be noted that the distribution center of 2D dispersion vector is assumed to be the origin under normal conditions in Fig. 7.32, which can be realized by transforming the coordinate systems. The missing alarm probability is expressed as

$$P_m = \iint_{\frac{x_1^2}{\sigma_1^2} + \frac{x_2^2}{\sigma_2^2} \leq d_{MT}^2} \frac{1}{2\pi\sigma_{A1}\sigma_{A2}} \exp\left[-\frac{1}{2}\left(\frac{(x_1 - \mu_{A1})^2}{\sigma_{A1}^2} + \frac{(x_2 - \mu_{A1})^2}{\sigma_{A1}^2}\right)\right] dx_1 dx_2 \tag{7.47}$$

The expression above is difficult to be analyzed. Generally, the anomaly of semi-major axis is independent of that of the inclination, thus capable of being separately analyzed. The probability of missing alarm is analyzed by a one-dimensional case below. As shown in Fig. 7.33, in normal conditions the dispersion value normally complies with the Gaussian distribution $x \sim N(0, \sigma)$ where the mean value is zero and the standard deviation is σ ; in abnormal conditions the dispersion value complies with the Gaussian distribution $x \sim N(\mu_A, \sigma_A)$ where the mean value is μ_A and the standard deviation is σ_A .

In the figure above, the threshold of Mahalanobis Distance is d_{MT} , the areas of two shaded parts are false alarm probability P_{fa} and missing alarm probability P_m , which can be separately expressed as follows:

Fig. 7.33 Probabilities of false and missing alarm (one-dimensional case)



$$\begin{aligned}
 P_{fa} &= 1 - \int_{-d_{MT}\sigma}^{+d_{MT}\sigma} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx = 2(1 - \Phi(d_{MT})) \\
 P_m &= \int_{-d_{MT}\sigma}^{+d_{MT}\sigma} \frac{1}{\sqrt{2\pi}\sigma_A} e^{-\frac{(x-\mu_A)^2}{2\sigma_A^2}} dx = \Phi\left[\left(d_{MT} - \frac{\mu_A}{\sigma}\right) \frac{\sigma}{\sigma_A}\right]
 \end{aligned}
 \tag{7.48}$$

The equations above can be used to describe the relationship between the threshold of Mahalanobis Distance, false alarm probability, and missing alarm probability. Generally, the threshold of Mahalanobis Distance is firstly determined by the first equation of (7.48) according to the false alarm probability needed and then the missing alarm probability is calculated by using information like the threshold of Mahalanobis Distance and orbital errors, etc.

7.5.4 Orbit Precision Requirement

The missing alarm probability is associated with σ/σ_A and μ_A/σ when the false alarm probability or the threshold of Mahalanobis Distance d_{MT} is definite. It decreases as σ/σ_A decreases and μ_A/σ increases. In the detection of orbital anomalies, σ means determined or predicted orbital errors, which represents orbital precision; μ_A refers to the variation scope of orbital parameters caused by orbital anomalies, which represents anomaly amplitude; and σ_A stands for the uncertainty in orbital parameter variation. Of the three parameters, μ_A and σ_A are decided by the orbital anomaly problem which needs to be handled. Their magnitudes are connected with the size and precision of orbital anomalies. σ is decided by the

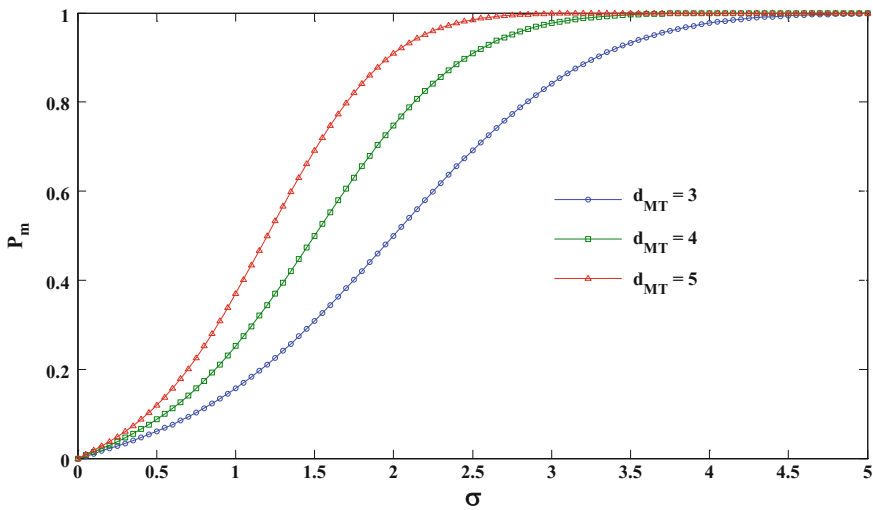


Fig. 7.34 Curve of missing alarm probability changing with orbital errors

orbital data which needs to use. The curve of false alarm probability changing with the orbital error σ is shown in Fig. 7.34 when $\mu_A = 6$ and $\sigma_A = 3$ and the threshold of Mahalanobis Distance is taken from different values.

As displayed in Fig. 7.34, the larger the threshold of Mahalanobis Distance is (namely the smaller the false alarm probability is), the smaller the missing alarm probability becomes, vice versa. When the threshold is fixed (namely the false alarm probability is certain), the missing alarm probability gradually increases to 1, as the orbital errors increase, vice versa.

When the mean value of dispersion and errors that are caused by orbital anomalies are known, if the threshold of Mahalanobis Distance and the acceptable missing alarm probability are given, the requirement on orbital precision can be calculated by Eq. (7.48). The requirement can be considered in two aspects. One is that the orbital error cannot be too big; otherwise, it will enhance misjudgment probability. The other is that small maneuvers cannot be detected (with respect to orbital errors).

References

1. Lemmens S, Krag H (2014) Two—line—elements—based maneuver detection methods for satellites in low earth orbit. *J Guid Control Dyn* 37(3):860–868
2. Patera RP (2008) Space event detection method. *J Spacecraft Rockets* 45(3):554–559
3. Kelecy T, Hall D, Hamada K, Stocker D (2007) Satellite maneuver detection using two—line element (TLE) data. Proceedings of the advanced Maui optical and space surveillance technologies conference, Maui, Hawaii
4. Hujsak R (2005) Reverse engineering maneuvers. In: AGI 2005 User's conference, Washington, DC
5. Flohrer T (2012) Classification of geosynchronous objects produced with the DISCOS database. ESA TR-14, Darmstadt, Germany
6. Kelecy T, Moriba J (2010) Detection and orbit determination of a satellite executing low thrust maneuvers. *Acta Astronaut* 66:798–809
7. Swartz RL, Coggi J, McNeill J (2010) A swift SIFT for satellite event detection. In: AIAA 2010-7527. AIAA/AAS Astrodynamics Specialist Conference, Toronto, Ontario Canada
8. McNeil JF, Coggi JM, Ailor WH et al (2009) Space situation monitoring laboratory: an integrated web—based environment for space environment information and analysis. In: AAS 09-416. Proceedings of the AAS/AIAA Astrodynamics Specialist Conference, Pittsburg, PA
9. Krag H, Klinkrad H, Flohrer T, Fletcher E (2010) The European surveillance and tracking system—services and design drivers. In: AIAA 2010-1927. SpaceOps 2010 Conference, Huntsville, Alabama
10. Sansgundo M, Molina MA (2010) Space debris in routine satellite operational; risk mitigation for collision. In: Proceedings of the SpaceOps 2010 conference, Huntsville, USA
11. Richmond D (2002) Space situational awareness research findings. In: AIAA/AAS Astrodynamics special conference, Monterey, CA
12. Kraus WF, Bowman C, George E et al (2011) Assessing abnormal space catalog updates. In: AIAA 2011-1627. proceedings of Info—tech@ Aerospace 2011 conference, St. Louis, Missouri
13. Dong YF, Su JM (2004) Detection of space target orbit maneuver on board by wavelet analysis. *J Astronautics* 25(2):213–218

14. Zhang ZJ, Dong YF (2008) Autonomic detection method of space target orbit maneuver based on wavelet analysis. *J Beijing Univ Aeronaut Astronautics* 34(3):353–356
15. Qiang S, Yi DY, Pan XG (2009) Research on space-based orbit maneuver identification algorithms. *J Syst Simul* 21(12):3528–3531
16. Yang X, Liu J, Wu XB et al (2009) New method to analyse the orbital abnormal of LEO Satellite using TLE Data-compositive criterion. *Space Debris Res Appl* (4):6–11
17. Wang RL, Gong JC, Zhang QM (2011) A simple efficient analytical method for orbital anomaly. *Space Debris Res Appl* 1:12–17
18. Song WD, Wang RL, Wang J (2012) A simple and valid analysis method for orbit anomaly detection. *Adv Space Res* (49):386–391
19. Yates JM, Spanbauer BW, Black JT (2010) Geostationary orbit development and evaluation for space situational awareness (GODESSA). In: *Proceedings of the AIAA/AAS Astrodynamics Specialist Conference*, Toronto, Ontario Canada
20. Aaron BS (2006) Geosynchronous satellite maneuver detection and orbit recovery using ground based optical tracking. Massachusetts Institute of Technology, Cambridge
21. Folcik ZJ, Cefola PJ, Abbot RI (2008) GEO maneuver detection for space situational awareness. *Adv Astronaut Sci* 129(1):523–550
22. Abbot RI, Wallace TP (2007) Decision support in space situational awareness. *Lincoln Lab J* 16:297–335
23. Hodge VJ, Austin J (2004) A survey of outlier detection methodologies. *Artif Intell Rev* 22:85–126
24. Ma ZH (2000) *Handbook of modern applied mathematics: volume for probability statistics and stochastic process*. Tsinghua University Press, Beijing
25. Zhang YT, Fang KT (1982) *Introduction of multivariate statistical analysis*. Science Press, Beijing
26. Brouwer D (1959) Solution of the problem of artificial satellite theory without drag. *Astron J* 64(1274):378–397
27. Kozai Y (1959) The motion of a close earth satellite. *Astron J* 64(1274):367–377
28. Hoots FR, Schumacher PW, Glover RA (2004) History of analytical orbit modeling in the U. S. space surveillance system. *J Guidance, Control Dyn* 27(2):174–185
29. Yang JC (1995) *Orbital dynamics and control of spacecraft: volume one*. China Astronautic Publishing House, Beijing
30. United States Strategic Command (USSTRATCOM). Space—track catalog number query results. [EB/OL]. [2012-04-16]. http://www.space-track.org/perl/id_query.pl
31. Soop EM (1999) *Handbook of geostationary orbits*. Wang, Z.C., Xing, G.H., Zhang, H.W., et al. National Defense Industry Press, Beijing
32. Agrawal BN (1992) *Design of geosynchronous spacecraft*. Chu, G.B., Zhang, J.M., Li, Y., et al. Aviation Industry Press, Beijing
33. Tombasco J, Axelrad P, Jah M (2010) Specialized coordinate representation for dynamic modeling and orbit estimation of geosynchronous orbits. *J Guid Control Dyn* 33(6): 1824–1836
34. Li HN (2010) *Geostationary satellites collocation*. National Defense Industry Press, Beijing

Appendix A

Elements of T-H Equation

State-Transition Matrix

Symbols in this book are defined as follows:

θ_0	True anomaly at the initial time
θ	True anomaly at a given time t
p	Semi-latus rectum of eccentric orbit, $p = h^2/\mu$
ρ	$\rho = 1 + e \cos \theta$
s	$s = \rho \sin \theta$
c	$c = \rho \cos \theta$
d	$d = -e \sin \theta$
ρ_0	$\rho_0 = 1 + e \cos \theta_0$
s_0	$s_0 = \rho_0 \sin \theta_0$
c_0	$c_0 = \rho_0 \cos \theta_0$
d_0	$d_0 = -e \sin \theta_0$
s'	$s' = \cos \theta + e \cos 2\theta$
c'	$c' = -(\sin \theta + e \sin 2\theta)$
k^2	$k^2 = h/p^2$
J	$J = k^2(t - t_0)$

The concrete expressions of all the elements of T-H equation state-transition matrix are provided below:

$$t_{11} = \left(\frac{3e\rho_0s(e^2 + 3\rho_0 - 1)}{\rho(e^2 - 1)} - \frac{3e^3s_0^2}{\rho\rho_0(e^2 - 1)} \right) J + \frac{es_0[2es_0 - cs_0 + s(c_0 - 2e)]}{\rho\rho_0(e^2 - 1)} + \frac{\rho_0[-2(e^2 + 3\rho_0 - 1) + 3c(e + c_0/\rho_0) + 3ss_0(1/\rho_0 + e^2/\rho_0^2)]}{\rho(e^2 - 1)}$$

$$t_{12} = \left(-\frac{3e^2\rho_0ss_0}{\rho(e^2 - 1)} \right) J - \frac{es_0[ce + (cc_0 + ss_0)(1/\rho_0 + 1) - 2\rho_0^2]}{\rho\rho_0(e^2 - 1)}$$

$$\begin{aligned}
t_{14} &= \left(\frac{3e^2 ss_0}{k^2 \rho \rho_0 (e^2 - 1)} \right) J + \frac{cs_0 - c_0 s + 2es - 2es_0}{k^2 \rho \rho_0 (e^2 - 1)} \\
t_{15} &= \left(\frac{3e \rho_0 s}{k^2 \rho (e^2 - 1)} \right) J + \frac{ce + (ss_0 + cc_0)(1/\rho_0 + 1) - 2\rho_0^2}{k^2 \rho \rho_0 (e^2 - 1)} \\
t_{21} &= \left(\frac{3\rho \rho_0 (e^2 + 3\rho_0 - 1)}{(e^2 - 1)} - \frac{3e^2 \rho s_0^2}{\rho_0 (e^2 - 1)} \right) J \\
&\quad + \frac{3\rho_0 [es_0(1/\rho_0 + 1/\rho_0^2) - s(1/\rho + 1)(e + c_0/\rho_0)]}{\rho (e^2 - 1)} \\
&\quad + \frac{3\rho_0 cs_0(1/\rho_0 + e^2/\rho_0^2)(1/\rho + 1)}{\rho (e^2 - 1)} \\
&\quad + \frac{es_0 [c_0 e - 2 + ss_0(1/\rho + 1) + c(1/\rho + 1)(c_0 - 2e)]}{\rho \rho_0 (e^2 - 1)} \\
t_{22} &= \left(-\frac{3e \rho \rho_0 s_0}{e^2 - 1} \right) J + \frac{\rho_0}{\rho} \\
&\quad - \frac{es_0 [es_0(1/\rho_0 + 1) - s(1/\rho + 1)[e + c_0(1/\rho_0 + 1)] + cs_0(1/\rho + 1)(1/\rho_0 + 1)]}{\rho \rho_0 (e^2 - 1)} \\
t_{24} &= \left(\frac{3e \rho s_0}{k^2 \rho_0 (e^2 - 1)} \right) J - \frac{c_0 e - 2 + ss_0(1/\rho + 1) + c(1/\rho + 1)(c_0 - 2e)}{k^2 \rho \rho_0 (e^2 - 1)} \\
t_{25} &= \left(\frac{3\rho \rho_0}{k^2 (e^2 - 1)} \right) J \\
&\quad + \frac{es_0(1/\rho_0 + 1) - s(1/\rho + 1)[e + c_0(1/\rho_0 + 1)] + cs_0(1/\rho + 1)(1/\rho_0 + 1)}{k^2 \rho \rho_0 (e^2 - 1)} \\
t_{33} &= [\rho_0 \cos(\theta - \theta_0) + d_0 \sin(\theta - \theta_0)]/\rho \\
t_{36} &= \sin(\theta - \theta_0)/(k^2 \rho \rho_0) \\
t_{41} &= \left\{ \frac{\rho_0 (3e^2 k^2 s^2 + 3s' e k^2 \rho^2)(e^2 + 3\rho_0 - 1)}{\rho (e^2 - 1)} - \frac{3e^4 k^2 s^2 s_0^2 + 3s' e^3 k^2 \rho^2 s_0^2}{\rho \rho_0 (e^2 - 1)} \right\} J \\
&\quad + \frac{3k^2 \rho_0 es(e^2 + 3\rho_0 - 1)}{\rho (e^2 - 1)} + \frac{3k^2 \rho \rho_0 [c'(e + c_0/\rho_0) + s's_0(1/\rho_0 + e^2/\rho_0^2)]}{e^2 - 1} \\
&\quad - \frac{es_0 k^2 \rho [c's_0 - s'(c_0 - 2e)]}{\rho_0 (e^2 - 1)} - \frac{3k^2 e^3 s_0^2}{\rho \rho_0 (e^2 - 1)} \\
&\quad + \frac{\rho_0 e k^2 s [3c(e + c_0/\rho_0) - 2(e^2 + 3\rho_0 - 1) + 3ss_0(1/\rho_0 + e^2/\rho_0^2)]}{\rho (e^2 - 1)}
\end{aligned}$$

$$\begin{aligned}
t_{42} &= \left(-\frac{3s'e^2k^2\rho\rho_0s_0}{e^2-1} - \frac{3e^3k^2\rho s^2s_0}{\rho(e^2-1)} \right) J - \frac{ek^2\rho s_0[c'e + (c'c_0 + s's_0)(1/\rho_0 + 1)]}{\rho_0(e^2-1)} \\
&\quad + \frac{e^2k^2ss_0[ce + (cc_0 + ss_0)(1/\rho_0 + 1) + \rho_0^2]}{\rho\rho_0(e^2-1)} \\
t_{44} &= \left(\frac{3e^3s^2s_0 + 3s'e^2\rho^2s_0}{\rho\rho_0(e^2-1)} \right) J + \frac{\rho(2s'e - c_0s' + c's_0)}{\rho_0(e^2-1)} + \frac{2e^2s^2 + s_0e^2s - c_0es^2 + cs_0es}{\rho\rho_0(e^2-1)} \\
t_{45} &= \left(\frac{3s'e\rho^2\rho_0 + 3e^2\rho_0s^2}{\rho(e^2-1)} \right) J + \frac{\rho\{c'[e + c_0(1/\rho_0 + 1)] + s's_0(1/\rho_0 + 1)\}}{\rho_0(e^2-1)} \\
&\quad + \frac{es\{c[e + c_0(1/\rho_0 + 1)] + \rho_0^2 + ss_0(1/\rho_0 + 1)\}}{\rho\rho_0(e^2-1)} \\
t_{51} &= \left\{ \frac{3e^3k^2\rho ss_0^2}{\rho_0(e^2-1)} - \frac{3ek^2\rho\rho_0s(e^2 + 3\rho_0 - 1)}{(e^2-1)} \right\} J - \frac{k^2\rho es_0[3es_0 - s_0(2c - e) + 2s(c_0 - 2e)]}{\rho_0(e^2-1)} \\
&\quad - ek^2s\rho_0 \left[\frac{3es_0(1/\rho_0 + 1/\rho_0^2) - 3s(1/\rho + 1)(e + c_0/\rho_0) + 3cs_0(1/\rho_0 + e^2/\rho_0^2)(1/\rho + 1)}{\rho(e^2-1)} \right] \\
&\quad - k^2\rho\rho_0 \left[\frac{-3(e^2 + 3\rho_0 - 1) + 3(e + c_0/\rho_0)(2c - e) + 6ss_0(1/\rho_0 + e^2/\rho_0^2)}{e^2-1} \right] \\
&\quad - \frac{e^2k^2ss_0[c_0e - 2 + ss_0(1/\rho + 1) + c(1/\rho + 1)(c_0 - 2e)]}{\rho\rho_0(e^2-1)} \\
t_{52} &= \left(\frac{3e^2k^2\rho\rho_0ss_0}{e^2-1} \right) J + \frac{k^2\rho es_0\{[e + c_0(1/\rho_0 + 1)](2c - e) - 3\rho_0^2 + 2ss_0(1/\rho_0 + 1)\}}{\rho_0(e^2-1)} + \frac{ek^2\rho_0s}{\rho} \\
&\quad - \frac{e^2k^2ss_0\{es_0(1/\rho_0 + 1) - s(1/\rho + 1)[e + c_0(1/\rho_0 + 1)] + cs_0(1/\rho + 1)(1/\rho_0 + 1)\}}{\rho\rho_0(e^2-1)} \\
t_{54} &= \left(-\frac{3e^2\rho ss_0}{\rho_0(e^2-1)} \right) J + \frac{\rho[3es_0 - s_0(2c - e) + 2s(c_0 - 2e)]}{\rho_0(e^2-1)} \\
&\quad - \frac{es[c_0e - 2 + ss_0(1/\rho + 1) + c(1/\rho + 1)(c_0 - 2e)]}{\rho\rho_0(e^2-1)} \\
t_{55} &= \left(-\frac{3e\rho\rho_0s}{e^2-1} \right) J - \frac{\rho\{[e + c_0(1/\rho_0 + 1)](2c - e) - 3\rho_0^2 + 2ss_0(1/\rho_0 + 1)\}}{\rho_0(e^2-1)} \\
&\quad + \frac{es\{es_0(1/\rho_0 + 1) - s(1/\rho + 1)[e + c_0(1/\rho_0 + 1)] + cs_0(1/\rho + 1)(1/\rho_0 + 1)\}}{\rho\rho_0(e^2-1)} \\
t_{63} &= d_0k^2[\rho \cos(\theta - \theta_0) - d \sin(\theta - \theta_0)] - \rho_0k^2[d \cos(\theta - \theta_0) + \rho \sin(\theta - \theta_0)]
\end{aligned}$$

Appendix B

Expression for Covariance Matrix Calculated with C-W Equation

The covariance matrix $\mathbf{P}(t)$ is a symmetrical matrix and the error along W direction is decoupled with the errors along R and S directions and can be written as follows:

$$\mathbf{P}(t) = \begin{bmatrix} \mathbf{P}_{rr}(t) & \mathbf{P}_{rv}(t) \\ \mathbf{P}_{rv}^T(t) & \mathbf{P}_{vv}(t) \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & 0 & p_{14} & p_{15} & 0 \\ p_{12} & p_{22} & 0 & p_{24} & p_{25} & 0 \\ 0 & 0 & p_{33} & 0 & 0 & p_{36} \\ p_{14} & p_{24} & 0 & p_{44} & p_{45} & 0 \\ p_{15} & p_{25} & 0 & p_{45} & p_{55} & 0 \\ 0 & 0 & p_{36} & 0 & 0 & p_{66} \end{bmatrix}$$

The non-zero elements of position covariance matrix $\mathbf{P}_{rr}(t)$ are as follows:

$$\begin{aligned} p_{11} &= (\cos nt - 2)^2 \sigma_{x0}^2 + \sin^2 nt \sigma_{y0}^2 + \frac{\sin^2 nt}{n^2} \sigma_{v_x,0}^2 + \frac{4(1 - \cos nt)^2}{n^2} \sigma_{v_y,0}^2 \\ p_{12} &= (2 - \cos nt)(2 \sin nt - 3nt) \sigma_{x0}^2 + \sin nt(2 \cos nt - 1) \sigma_{y0}^2 \\ &\quad - \frac{2}{n^2} (1 - \cos nt) \sin nt \sigma_{v_x,0}^2 + \frac{2}{n} (\cos nt - 1) \left(3t - \frac{4}{n} \sin nt \right) \sigma_{v_y,0}^2 \\ p_{22} &= (2 \sin nt - 3nt)^2 \sigma_{x0}^2 + (2 \cos nt - 1)^2 \sigma_{y0}^2 \\ &\quad + \frac{4}{n^2} (1 - \cos nt)^2 \sigma_{v_x,0}^2 - \left(\frac{4}{n} \sin nt - 3t \right)^2 \sigma_{v_y,0}^2 \\ p_{33} &= \cos^2 nt \sigma_{z0}^2 + \frac{1}{n^2} \sin^2 nt \sigma_{v_z,0}^2 \end{aligned}$$

All the elements of the cross-correlated covariance matrix $\mathbf{P}_{rv}(t)$ between position and velocity are listed below:

$$\begin{aligned}
p_{14} &= n(\cos nt - 2)(3nt - \sin nt)\sigma_{x0}^2 + n \sin nt(\cos nt - 1)\sigma_{y0}^2 \\
&\quad - \frac{1}{n}\sin nt(2 - \cos nt)\sigma_{v_x0}^2 + \frac{2}{n}(3nt - 2 \sin nt)(1 - \cos nt)\sigma_{v_y0}^2 \\
p_{15} &= n(\cos nt - 2)(\cos nt - 1)\sigma_{x0}^2 - n \sin^2 nt \sigma_{y0}^2 - \frac{1}{n} \sin^2 nt \sigma_{v_x0}^2 \\
&\quad + \frac{2}{n}(1 - \cos nt)(2 \cos nt - 1)\sigma_{v_y0}^2 \\
p_{24} &= -n(2 \sin nt - 3nt)(\sin nt - 3nt)\sigma_{x0}^2 - n(2 \cos nt - 1)(\cos nt - 1)\sigma_{y0}^2 \\
&\quad - \frac{2}{n}(1 - \cos nt)(2 - \cos nt)\sigma_{v_x0}^2 + \left(\frac{4}{n}\sin nt - 3t\right)(3nt - 2 \sin nt)\sigma_{v_y0}^2 \\
p_{25} &= n(2 \sin nt - 3nt)(\cos nt - 1)\sigma_{x0}^2 - n \sin nt(2 \cos nt - 1)\sigma_{y0}^2 \\
&\quad + \frac{2}{n}(1 - \cos nt) \sin nt \sigma_{v_x0}^2 + \left(\frac{4}{n}\sin nt - 3t\right)(2 \cos nt - 1)\sigma_{v_y0}^2 \\
p_{36} &= -n \sin nt \cos nt \sigma_{z0}^2 + \frac{1}{n}\sin nt \cos nt \sigma_{v_z0}^2
\end{aligned}$$

The non-zero elements of velocity covariance matrix $\mathbf{P}_{vv}(t)$ are as follows:

$$\begin{aligned}
p_{44} &= n^2(\sin nt - 3nt)^2\sigma_{x0}^2 + n^2(\cos nt - 1)^2\sigma_{y0}^2 + (2 - \cos nt)^2\sigma_{v_x0}^2 \\
&\quad + (3nt - 2 \sin nt)^2\sigma_{v_y0}^2 \\
p_{45} &= n^2(\sin nt - 3nt)(1 - \cos nt)\sigma_{x0}^2 + n^2 \sin nt(\cos nt - 1)\sigma_{y0}^2 \\
&\quad - \sin nt(2 - \cos nt)\sigma_{v_x0}^2 + (3nt - 2 \sin nt)(2 \cos nt - 1)\sigma_{v_y0}^2 \\
p_{55} &= n^2(1 - \cos nt)^2\sigma_{x0}^2 + n^2 \sin^2 nt \sigma_{y0}^2 + \sin^2 nt \sigma_{v_x0}^2 + (2 \cos nt - 1)^2\sigma_{v_y0}^2 \\
p_{66} &= n^2 \sin^2 nt \sigma_{z0}^2 + \cos^2 nt \sigma_{v_z0}^2.
\end{aligned}$$

Appendix C

Some Figures of Chapter 3

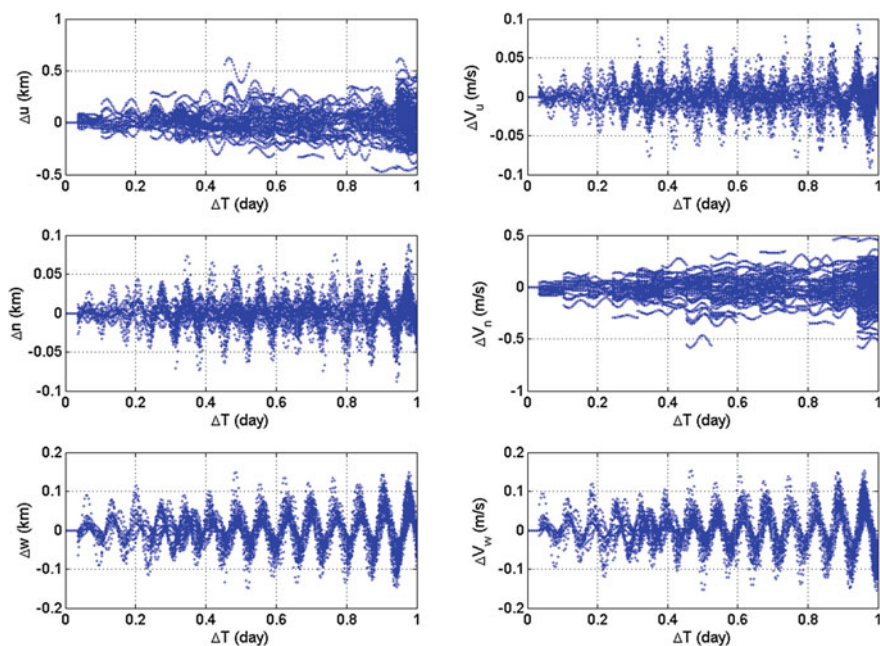


Fig. C.1 TLE's prediction residual state components in UNW coordinate system versus temporal difference for LEO object

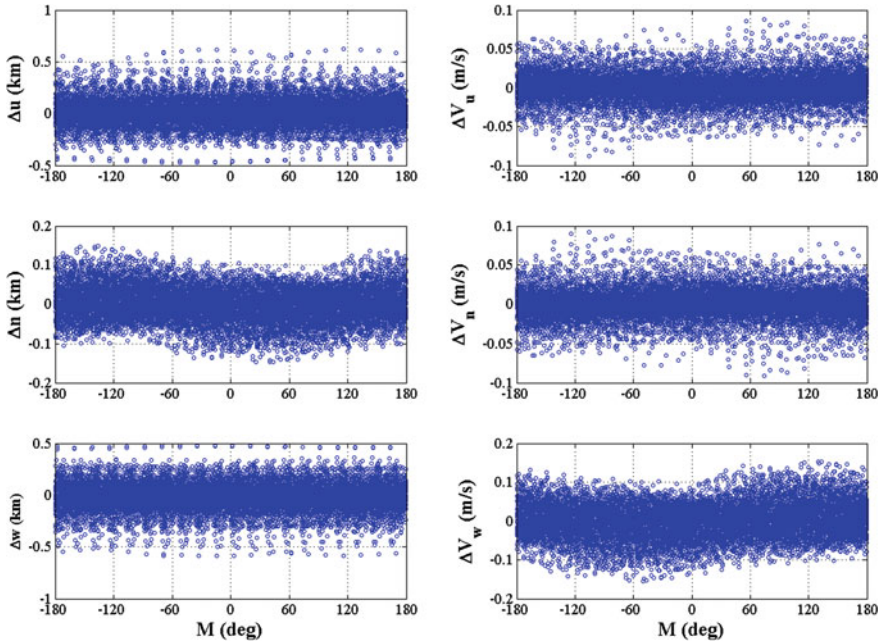


Fig. C.2 TLE's prediction residual state components in UNW coordinate system versus mean anomaly for LEO object

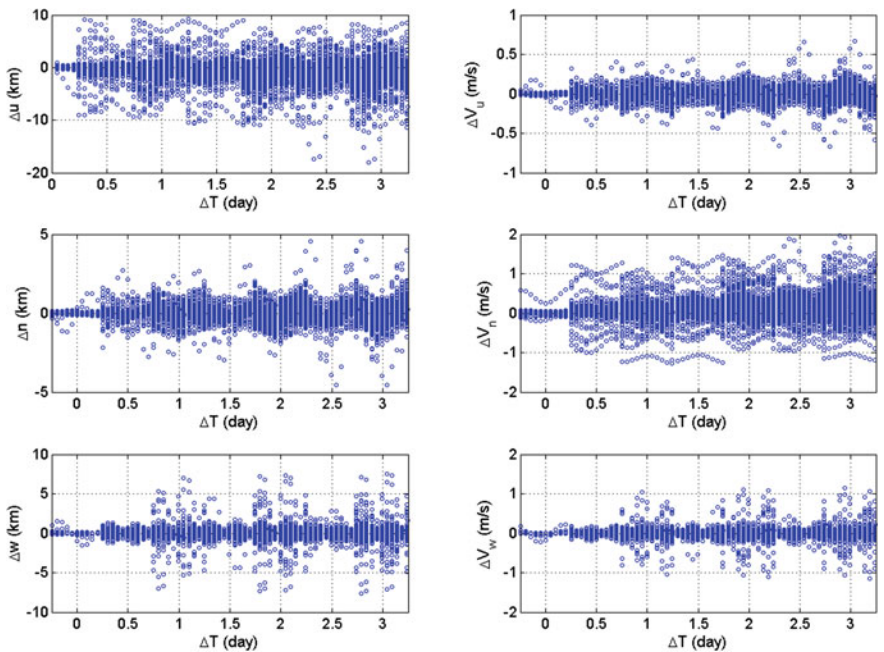


Fig. C.3 TLE's prediction residual state components in UNW coordinate system versus temporal difference for MEO object

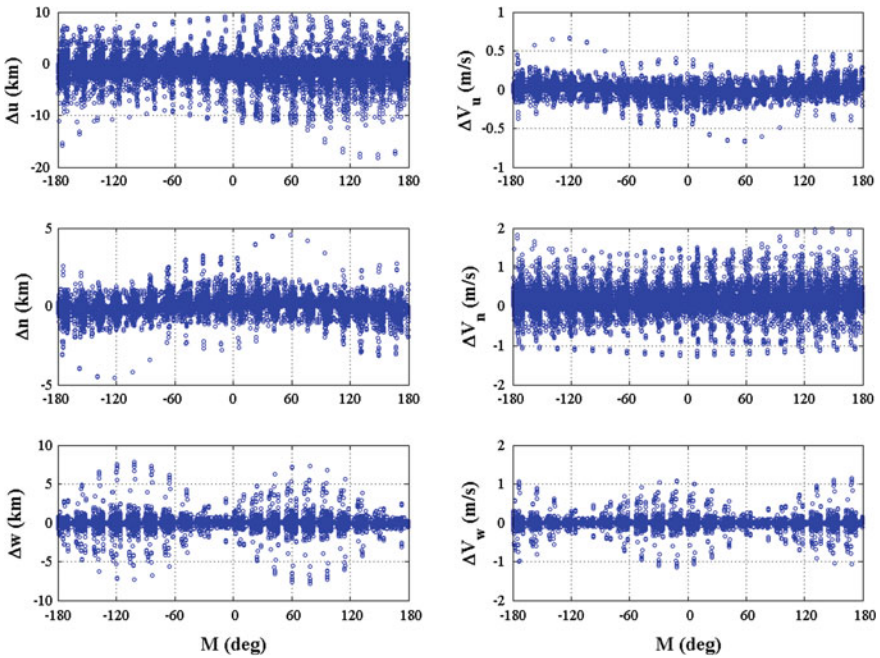


Fig. C.4 TLE’s prediction residual state components in UNW coordinate system versus mean anomaly for MEO object

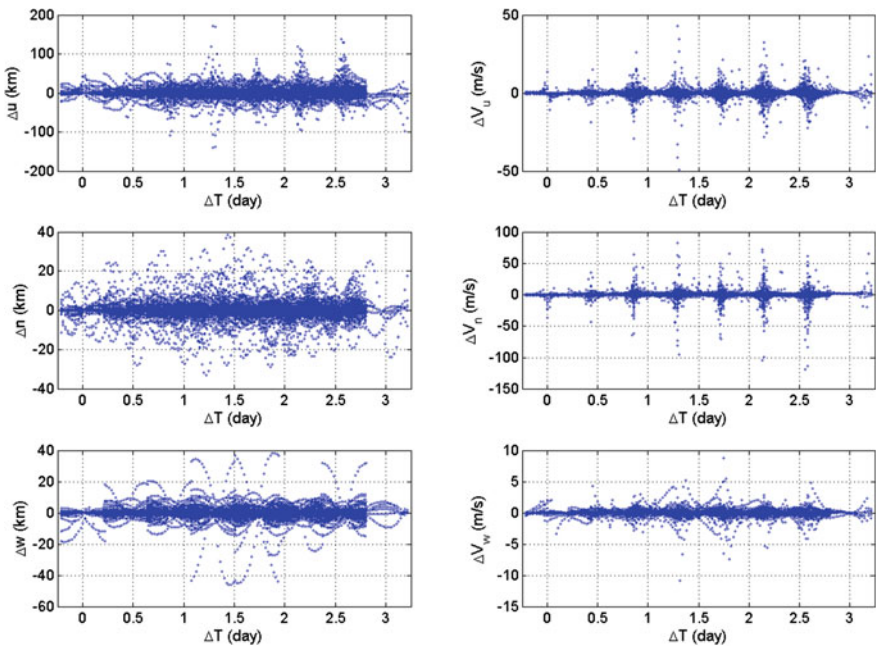


Fig. C.5 TLE’s prediction residual state components in UNW coordinate system versus temporal difference for HEO object

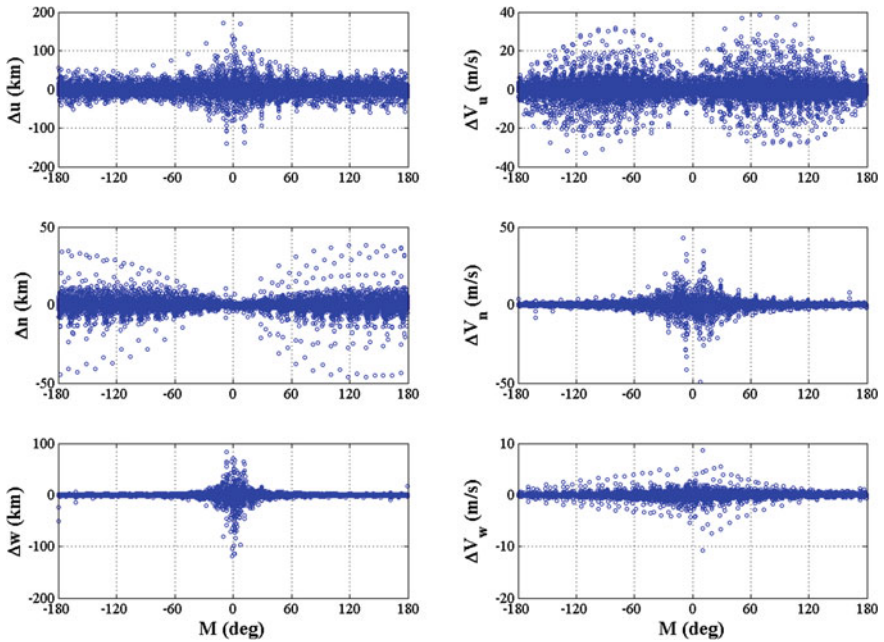


Fig. C.6 TLE's prediction residual state components in UNW coordinate system versus mean anomaly for HEO object

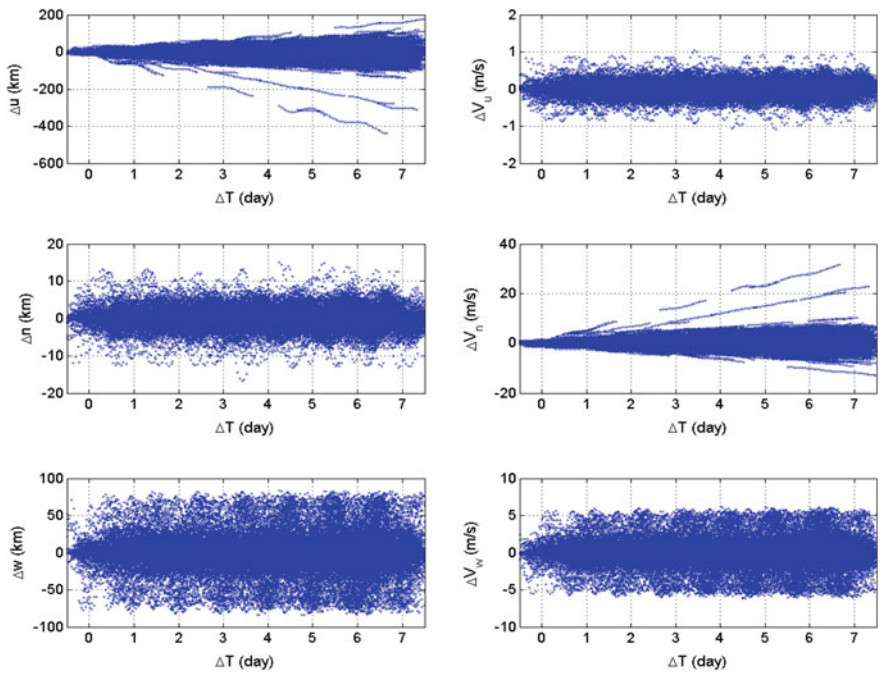


Fig. C.7 TLE's prediction residual state components in UNW coordinate system versus temporal difference for GEO object

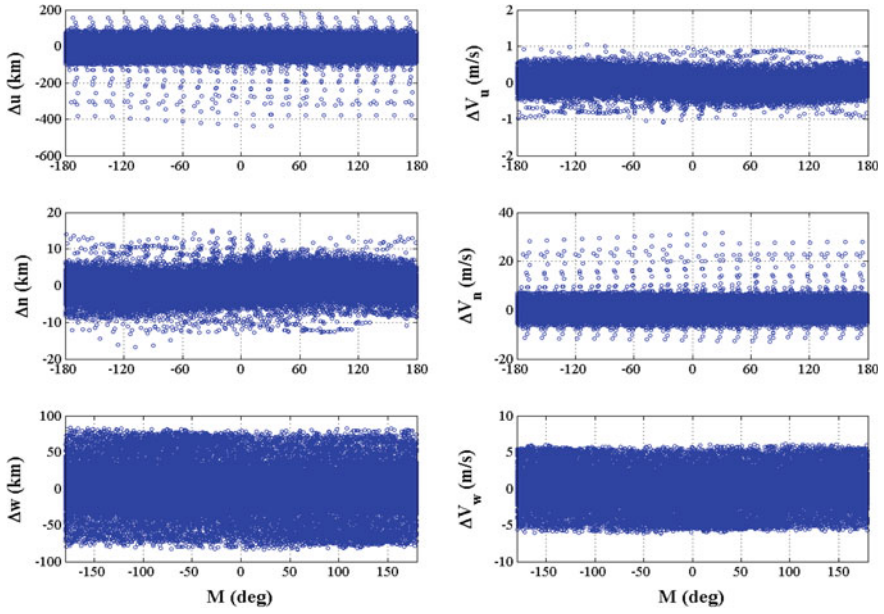


Fig. C.8 TLE's prediction residual state components in UNW coordinate system versus mean anomaly for GEO object

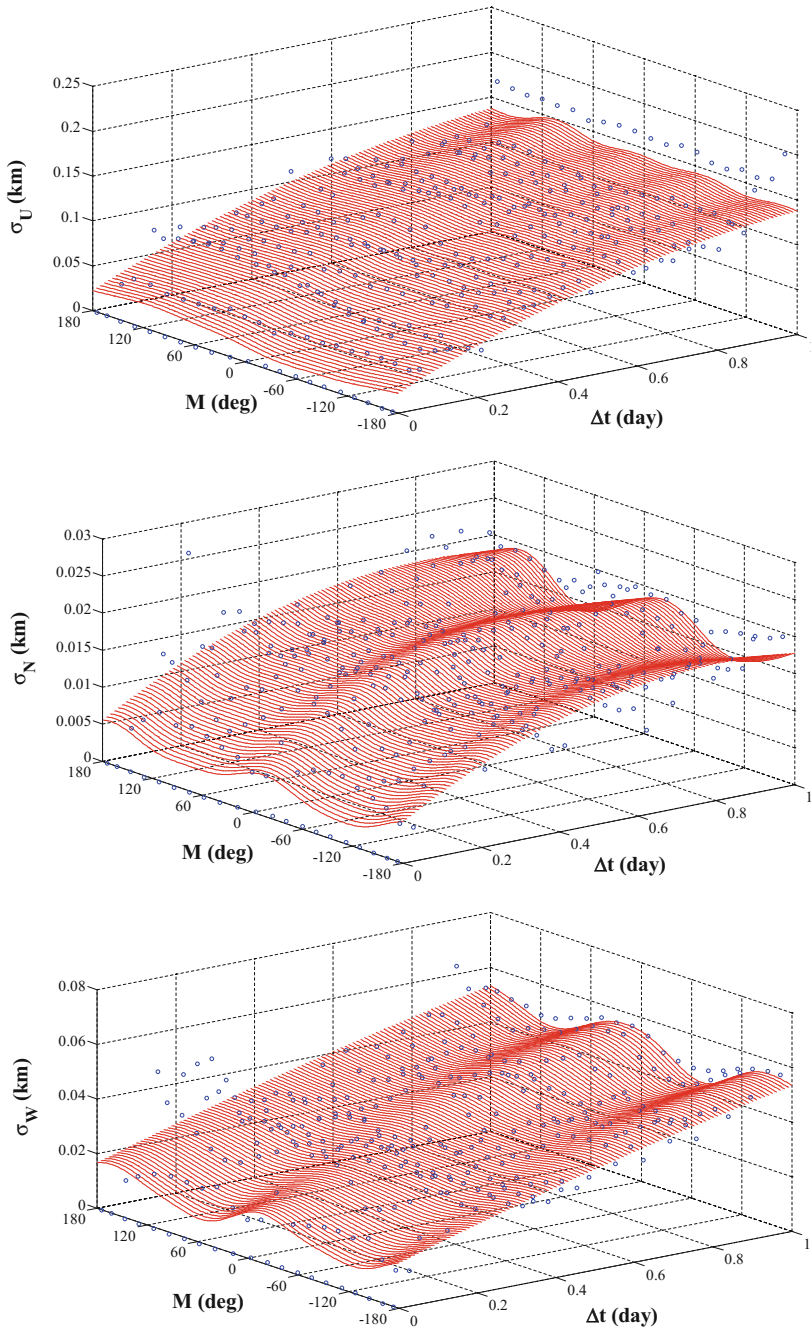


Fig. C.9 In-track (U), normal (N), and cross-track (W) error standard deviation's Poisson series fitting curved surfaces of LEO objects ($n = 2$, $m = 4$)

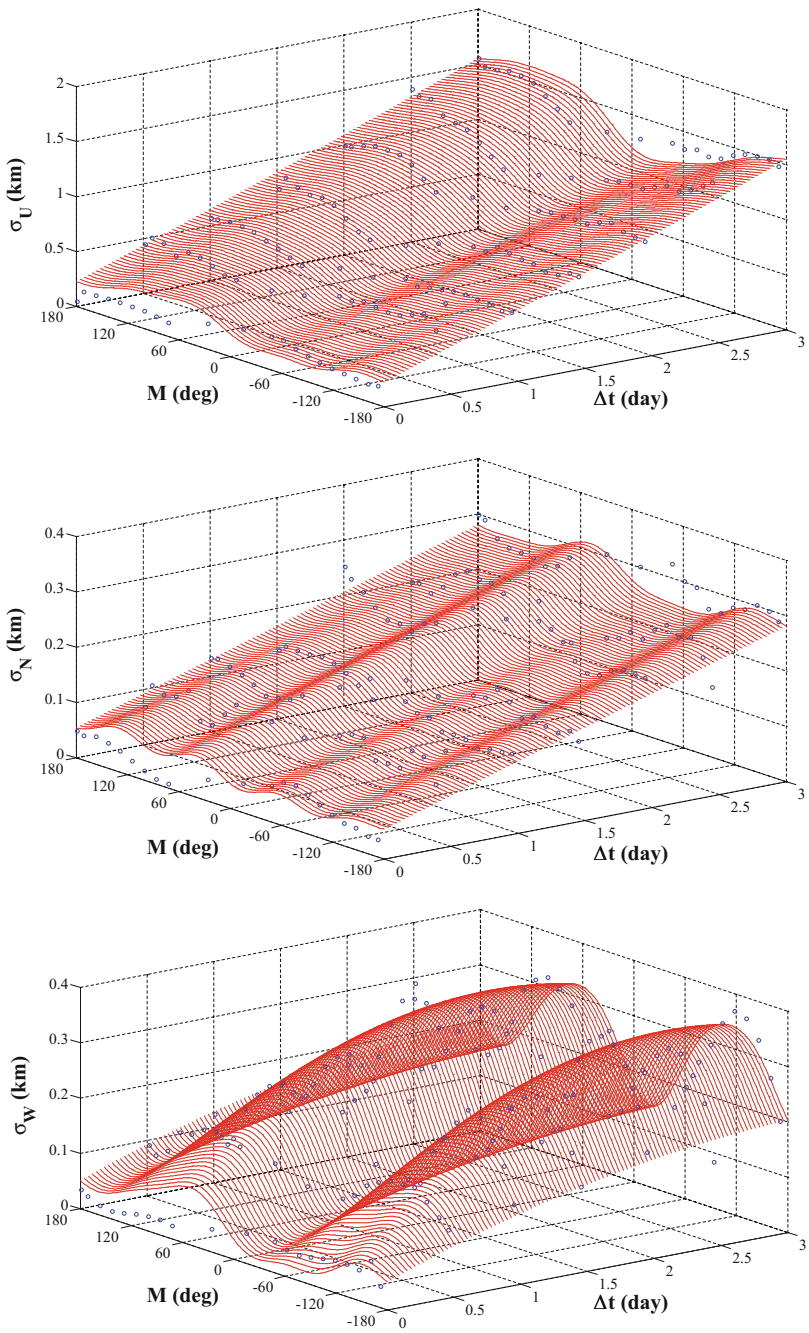


Fig. C.10 In-track (U), normal (N), and cross-track (W) error standard deviation's Poisson series fitting curved surfaces of MEO objects ($n = 2, m = 4$)

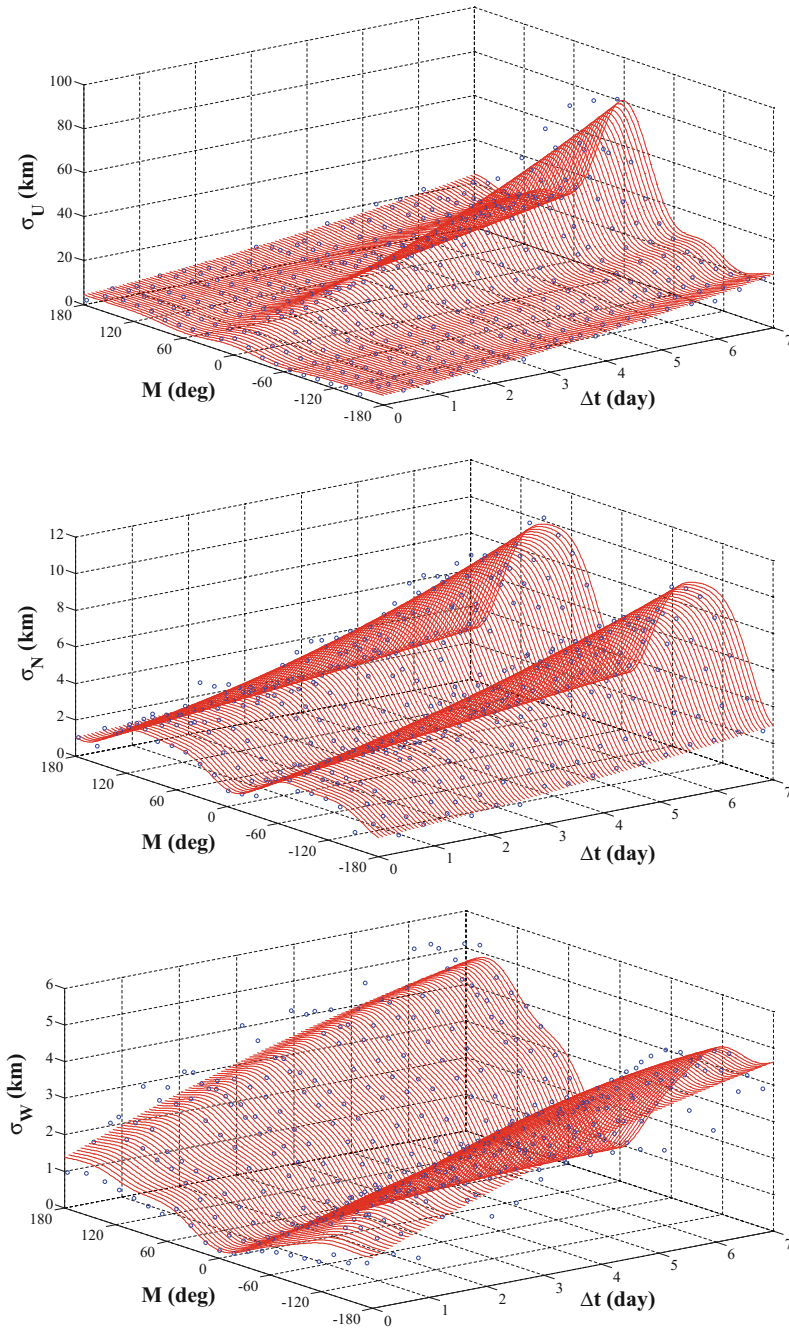


Fig. C.11 In-track (U), normal (N), and cross-track (W) error standard deviation's Poisson series fitting curved surfaces of HEO objects ($n = 2$, $m = 4$)

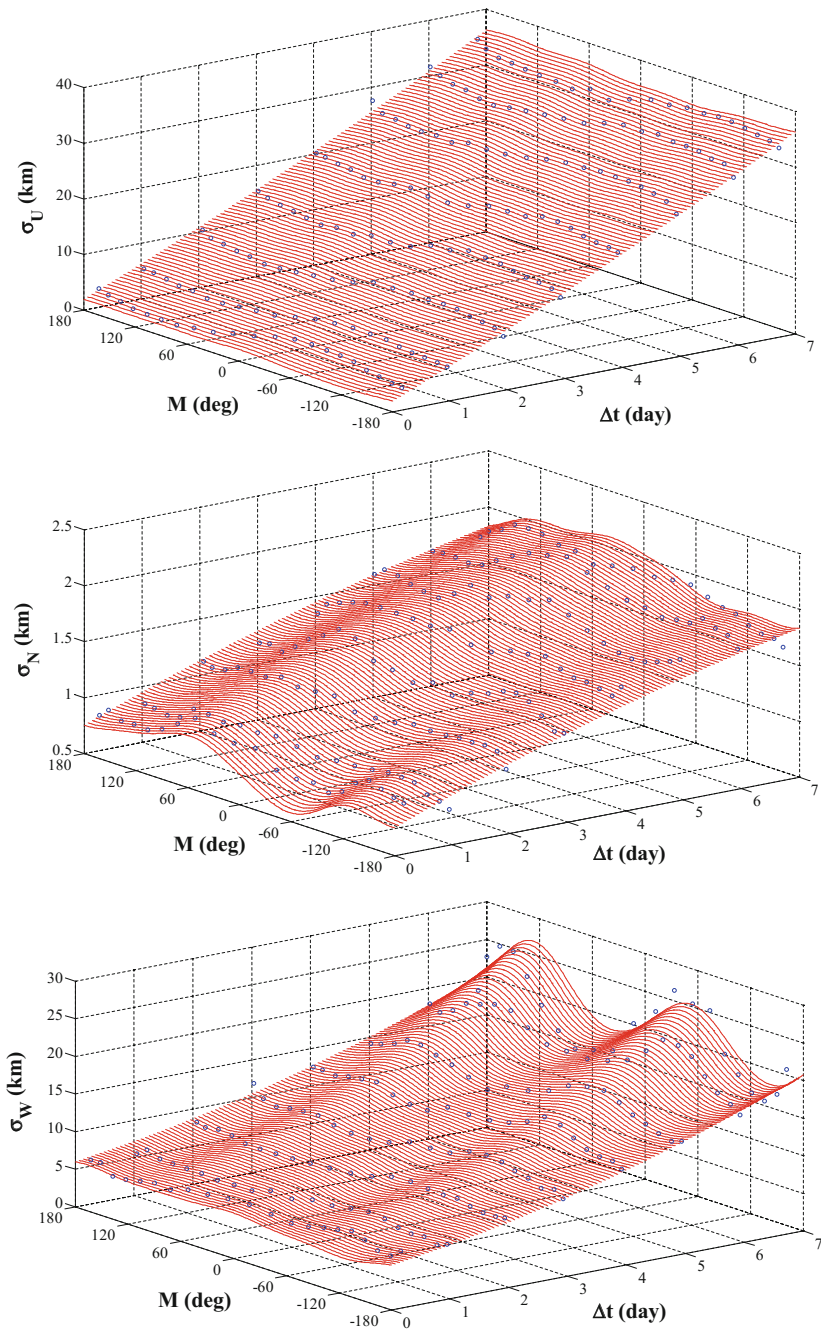


Fig. C.12 In-track (U), normal (N), and cross-track (W) error standard deviation's Poisson series fitting curved surfaces of GEO objects ($n = 2, m = 4$)

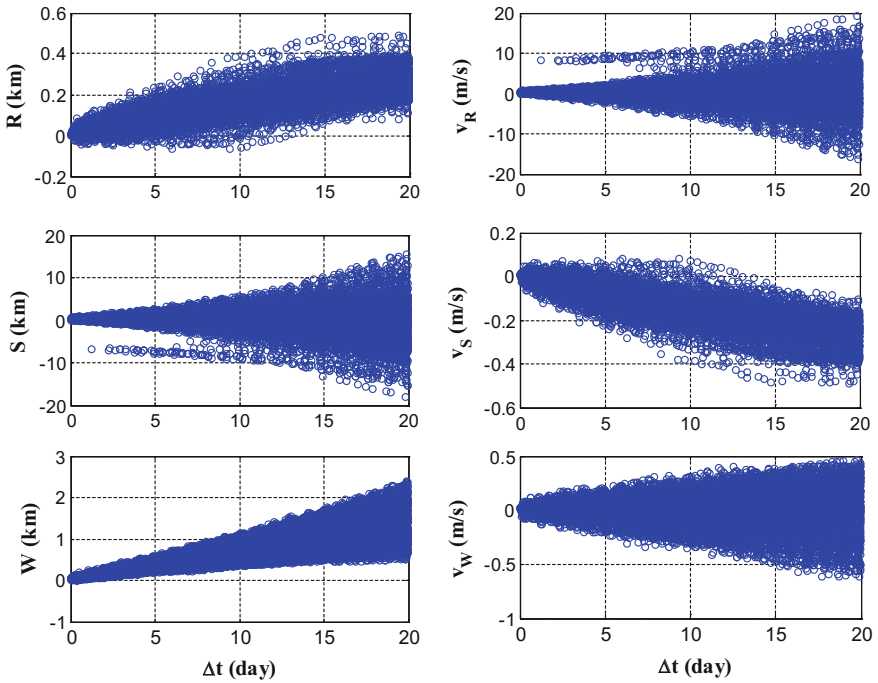


Fig. C.13 TLE's prediction residual state components in RSW coordinate system versus temporal difference for LEO object (20 days)

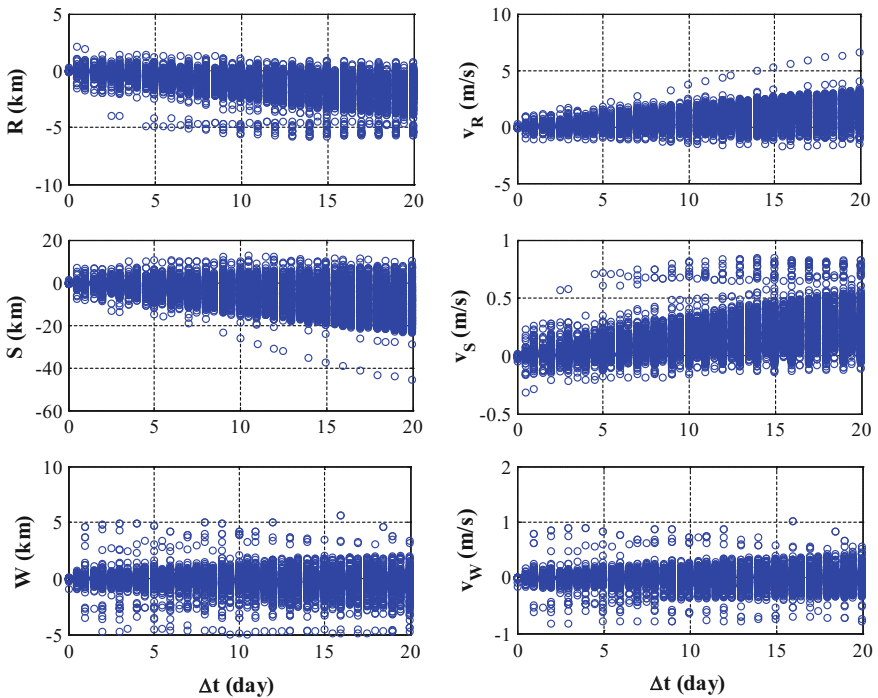


Fig. C.14 TLE's prediction residual state components in RSW coordinate system versus temporal difference for MEO object (20 days)

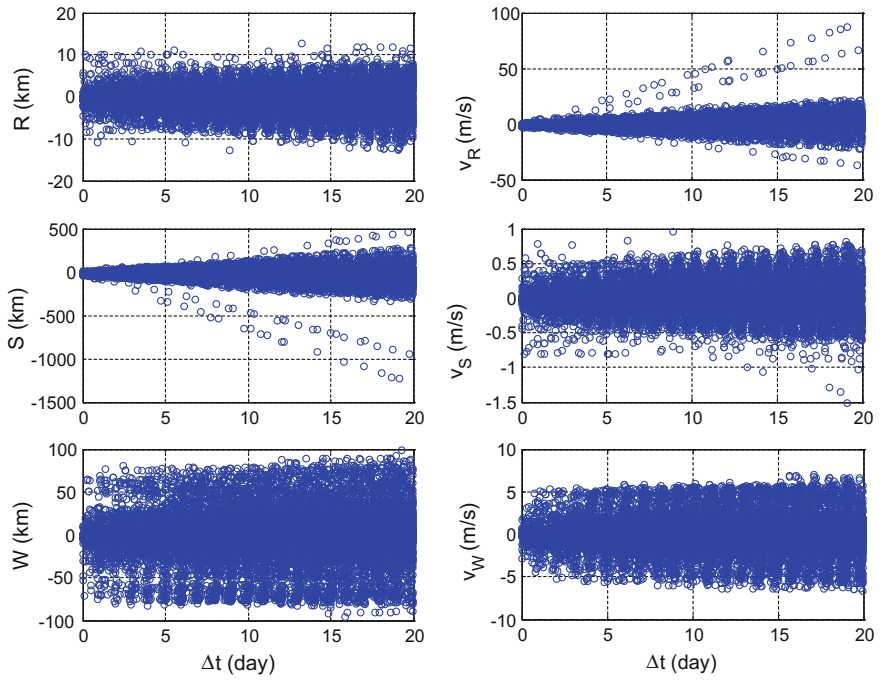


Fig. C.15 TLE’s prediction residual state components in RSW coordinate system versus temporal difference for GEO object (20 days)

Bibliography

1. Silver N (2012) *The signal and the noise*. Penguin Press
2. Zhu YL (1999) Environment model of space debris. *Space International* 3:13–16
3. Klindrad H (2006) *Space debris—models and risk analysis*. Springer
4. Duncan M, Long A (2006) Realistic covariance prediction for the Earth science constellation. AIAA 2006-6293-AIAA/AAS astrodynamics specialist conference and exhibit, Keystone, Colorado
5. Bird D (2013) Conjunction summary message guide. 64th international astronomical congress, Beijing, China
6. Qiu YM (2011) Research on control strategy of close-range rendezvous based on trajectory safety. National University of Defense Technology, Changsha
7. Zhu YL (1997) The latest situation of space debris environment. *Aerospace China* 1:22–28
8. Du D, Liu J (2002) Manned spaceflight and space debris. *Aerospace China* 2:18–23.
9. Xue FX, Yang XY (2004) Research overview of space debris. *Space International* 5:14–19
10. Zhu YL (1999) NASA's space debris models. *Aerospace Shanghai* 3:24–30
11. Kessler DJ, Matney MJ, Zhang J et al (1996) A computer based orbital debris environment model for spacecraft design and observation in low Earth orbit. TM-104825, 11
12. Zhang J, Kessler DJ, Matney MJ et al (1997) The NASA engineering model: a new approach. *Adv Space Res* 19(2):281–290
13. Liou JC, Matney M, Anz-Meador PD et al (2002) The new NASA orbital debris engineering model OR-DEM2000. NASA/TP-2002-210780
14. Xu YL, Horstman M, Krisko PH et al (2009) Modeling of LEO orbital debris populations for ORDEM2008. *Adv Space Res* 43:769–782
15. Krisko PH (2010) NASA's new orbital debris engineering model, ORDEM2010. 4th international association for the advancement of space safety, Huntsville, Alabama, 19–21 May 2010.
16. Matney M (2010) An overview of NASA's orbital debris environment model. *Adv Aeronaut Sci* 137:31–36
17. Krisko PH, Hall DT (2004) Geosynchronous region orbital debris modeling with GEO-EVOLVE 2.0. *Adv Space Res* 34:1166–1170
18. Reynolds R, Eichler P, Bade A et al (2009) Sensitivity analysis of the orbital debris environment using the EVOLVE 4.0 model. *Adv Space Res* 23(1):175–185
19. Klindrad H, Bendisch J, Bunte KD et al (2001) The MASTER-99 space debris and meteoroid environment model. *Adv Space Res* 28(9):1355–1366
20. Klindrad H, Sdunnus H (1997) Concepts and applications of the MASTER space debris environment model. *Adv Space Res* 19(2):277–280
21. Klindrad H, Sdunnus H, Bendisch J (1995) Development status of the ESA space debris reference model. *Adv Space Res* 16(11):93–102

22. Kanemitsu Y, Akahoshi Y, Narumi T, et al (2012) Comparison of space debris environment models: OR-DEM2000, MASTER-2001, MASTER-2065 and MASTER-2009. JAXA Research and Development Memorandum, JAXA-RM-11-020E.
23. Bendisch J, Bunte K, Klinkrad H et al (2004) The MASTER-2001 model. *Adv Space Res* 34:959–968
24. Flegel SK, Celhaus J, Mockel M et al (2011) Multi-layer insulation model for MASTER-2009. *Acta Astronautica* 69:911–922
25. Ye QX, Shen YH (2006) *Practical handbook of mathematics*-edition two. Science Press, Beijing
26. Xu K, Pang BJ (2008) Study on spatial density and flux algorithms of space debris, vol. 3. *Space Debris Research*
27. Tang Q, Pang BJ, Zhang W (2005) Analysis of the parameters in space debris environment engineering model. *Chin Space Sci Technol* 204, 5: 22-27.
28. Li CA, Pang BJ (2008) Research on the spatial density algorithm of space debris. *Chin J Space Sci* 28(6):522–530
29. Cao YH (2013) Space debris environment modeling research based on historical orbit data. National University of Defense Technology, Changsha
30. Hill GW (1878) Researches in Lunar theory. *Am J Math* 1:5–26
31. Clohessy WH, Wiltshire RS (1960) Terminal guidance system for satellite rendezvous. *J Aerosp Sci* 27(9):653–674