

Nonlinear Filtering for Autonomous Navigation of Spacecraft in Highly Elliptical Orbit

by

Adam C. Vigneron, B.Sc.(Eng.)

A thesis submitted to the
Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree of

Master of Applied Science

Ottawa-Carleton Institute for Mechanical and Aerospace Engineering
Department of Mechanical and Aerospace Engineering
Carleton University
Ottawa, Ontario
May, 2014

©Copyright
Adam C. Vigneron, 2014

The undersigned hereby recommends to the
Faculty of Graduate and Postdoctoral Affairs
acceptance of the thesis

Nonlinear Filtering for Autonomous Navigation of Spacecraft in Highly Elliptical Orbit

submitted by **Adam C. Vigneron, B.Sc.(Eng.)**

in partial fulfillment of the requirements for the degree of

Master of Applied Science

Professor Anton H. J. de Ruiter, Thesis Supervisor

Professor Bruce Burlton, Thesis Co-supervisor

Professor Alex Ellery, Thesis Co-supervisor

Professor Metin Yaras, Chair,
Department of Mechanical and Aerospace Engineering

Ottawa-Carleton Institute for Mechanical and Aerospace Engineering
Department of Mechanical and Aerospace Engineering
Carleton University
May, 2014

Abstract

To fill a gap in satellite services for the Canadian Arctic, the Canadian Space Agency has proposed a Polar Communication and Weather (PCW) mission to be flown in a highly elliptical Molniya orbit. In an era of increasingly capable space hardware, autonomous satellite navigation has become a standard means by which satellites in low Earth orbit can increase their independence and functionality. This study examined the accuracy to which autonomous navigation might be realized in a Molniya orbit.

Using appropriate physical force models and simulated pseudorange signals from the Global Positioning System (GPS), a navigation algorithm based on the Extended Kalman Filter was demonstrated to achieve a three-dimensional root-mean-square accuracy of 58.9 m over a $500\text{ km} \times 40\,000\text{ km}$ Molniya orbit. This accuracy satisfied the requirements of the PCW mission and demonstrated the utility of GPS signal reception at high altitudes. Algorithms based on the Unscented Kalman Filter and the Cubature Kalman Filter were not found to improve this result; this was due to a high frequency of measurements during periods of highly nonlinear dynamics.

During this study, detailed models were developed for GPS pseudorange errors, including ephemeris errors, transmitter clock errors, and ionospheric delay. Receiver clock bias error was shown to be a significant source of navigation solution error; for reasons of geometry, the navigation algorithm is not able to differentiate between this error and a radial position error. GPS sidelobe signals were shown to be an effective means to acquire additional GPS signals over the highly elliptical orbit. Finally, an exploratory study found ground-based radio beacons to be a useful navigation aid for this orbital regime.

To Mom and Dad

Acknowledgments

As one chapter in life ends and another begins, a few words of thanks are in order:

To my family, in Saskatchewan and in Ottawa; you are the rock upon which these last three years have been built. Thank you for the words of wisdom, for the warm meals, and for always being just a phone call away.

To my friends; through sport and study, through activities social and spiritual, you kept my keel even and my goals in sight. Thank you for filling these years in Ottawa with light, laughter, and song.

To Warren and the rest of the staff at Magellan; it was a privilege to be taken under your collective wing. Thank you for the unique and wonderful opportunity to become your colleague and friend.

To my co-supervisor, Prof. Bruce Burlton; thank you for bringing experience, good humour, and direct advice to the many ups and downs of this endeavour.

Above all, to my supervisor, Dr. Anton de Ruiter; your patience, trust, and persistence are beyond my words to describe. For the early emails and the late-night phone calls, for the freedom to define my own pace, and for unwavering dedication to this project: thank you.

Adam C. Vigneron

16 May 2014

Ottawa, Canada

Table of Contents

Abstract	iii
Acknowledgments	v
Table of Contents	vi
List of Tables	x
List of Figures	xi
Nomenclature	xiii
1 Introduction	1
1.1 Motivation for this study: Satellite Service in Canada's North	1
1.2 Polar Communication and Weather	3
1.2.1 The Molniya Orbit	3
1.3 The Need for Navigation	4
1.3.1 The Next Step: Autonomous Navigation	6
1.3.2 Autonomous Navigation in HEO	7
1.3.3 The Magellan Study on HEO Navigation	8
1.4 Problem Definition: Navigation on HEO	8
1.5 Three Keys to Navigation	9
1.5.1 Predicting: refining our expectations	10
1.5.2 Measuring: selecting our observations	11
1.5.3 Filtering: fusing for navigation	11
1.6 Organization of this thesis	12

2 Physical Models and Propagation	13
2.1 Introduction	13
2.1.1 Dead Reckoning: The Art of Propagation	13
2.1.2 Accelerations at play on the HEO orbit	15
2.2 Perturbing Accelerations and Their Models	16
2.2.1 Earth Gravity	16
2.2.2 Third-Body Gravity	19
2.2.3 Atmospheric Drag	20
2.2.4 Solar Radiation Pressure	22
2.2.5 Secondary Perturbations	23
2.3 Verifying the Propagator	26
2.3.1 Impact of the Perturbing Accelerations	27
2.3.2 Propagator Accuracy	29
2.4 Chapter Summary	30
2.5 Propagator Configuration Tables	31
3 Measurements and their Models	33
3.1 Introduction	33
3.2 Measurements for Autonomous Navigation	34
3.3 The Global Positioning System: an introduction	36
3.3.1 A brief history of American radionavigation systems	36
3.3.2 GPS: Basis of Operation	36
3.3.3 GPS: System Architecture	37
3.4 GPS in this thesis: Models and Errors	38
3.4.1 Tropospheric Error	40
3.4.2 Ionospheric Error	40
3.4.3 SV Ephemeris Errors	42
3.4.4 SV Clock Errors	43
3.5 GPS in this thesis: Receiver Clock Error	44
3.5.1 Propagating the Receiver Clock	45
3.6 GPS in this thesis: Link Budget	48
3.6.1 Calculating the link budget: the GPS simulator	51
3.6.2 Applying the link budget: satellite visibility	52
3.6.3 Refining the link budget: side-lobes	55
3.7 Passive Ranging	56

3.8	Other Measurements	58
3.9	Chapter Summary	59
4	Nonlinear Filter Theory	61
4.1	The Discrete Linear Kalman Filter	62
4.1.1	Problem Definition	62
4.1.2	Algorithm	63
4.2	Nonlinear Kalman Filters	64
4.2.1	Extended Kalman Filter	65
4.2.2	Unscented Kalman Filter	68
4.2.3	Cubature Kalman Filter	72
4.2.4	Continuous-Discrete Kalman Filtering	73
4.3	Kalman Flight Heritage	75
4.4	Chapter Summary	77
5	Nonlinear Filter Performance	78
5.1	Simulation Implementation	78
5.2	Preamble: Setting up the Simulation	79
5.2.1	Simulation Configuration	79
5.2.2	Propagator Configuration	80
5.2.3	Measurement Configuration	82
5.3	Baseline EKF Analysis	83
5.3.1	What do these figures mean?	83
5.3.2	Baseline EKF Performance	89
5.4	Perturbing the Baseline EKF	90
5.4.1	Effect of Receiver Antenna Positioning	90
5.4.2	Effect of EKF Propagator Fidelity	93
5.4.3	Effect of Sidelobes	95
5.4.4	Effect of Receiver Clock Error	97
5.4.5	Inclusion of Passive Ranging	99
5.5	Other Nonlinear Filters	102
5.6	Chapter Summary	105
6	Conclusions	107
6.1	Conclusions and Discussion	108

6.2	Contributions	109
6.3	Future Work	110
List of References		112
Appendix A Supplementary Mathematics		123
A.1	The Classical Orbital Elements	123
A.2	Reference Frames	124
A.2.1	The JPL Development Ephemeris	125
A.3	Time Dilution of Precision	126
A.3.1	TDOP for a Molniya orbit	127
A.3.2	Selecting a TDOP threshold	128
Appendix B Recursive Algorithm for Earth Gravity		129
Appendix C Simulation Details		131
C.1	Using the GPS simulator	131
C.2	The MATLAB simulation environment	133
Appendix D Solar Radiation Pressure: Derivation and Justification		134
Appendix E GPS Broadcast Errors		136

List of Tables

2.1	Gravitational Influence of Celestial Bodies at Earth	20
2.2	Radiation Pressure Coefficient	24
2.3	Initial Conditions for Molniya Orbit	25
2.4	Truth Orbits for Propagator Verification	27
2.5	Propagator Verification Results	29
2.6	Summary of propagator models	30
2.7	Propagator Constants and Parameters	31
2.8	STK 9.2.0 HPOP Force Models: Primary Perturbations	32
2.9	STK 9.2.0 HPOP Force Models: Secondary Perturbations	32
3.1	Steered receiver clock parameters	47
3.2	Drifting receiver clock parameters	47
3.3	Link Budget Constants	49
3.4	Summary of measurement errors and models for GPS pseudorange . .	60
5.1	Baseline Simulation Configuration	84
5.2	Tuning parameters for the Unscented Kalman Filter	102
5.3	Summary of Results for Nonlinear Filters	104
5.4	Filter Design: Decisions and Observations	106
E.1	SV ephemeris errors (RMS) for the day of 04 Feb 2007.	137
E.2	SV clock-correction errors (RMS) for the day of 04 Feb 2007.	138

List of Figures

1.1	Weather from GEO	2
1.2	Weather from LEO	2
1.3	A typical Molniya orbit.	5
1.4	A typical Molniya ground track.	5
2.1	Chip Log	14
2.2	Secondary Perturbations	25
2.3	Primary Perturbations	28
2.4	Propagator Verification	28
3.1	Families of hyperbolae from LORAN transmitters.	37
3.2	Antenna Gains	50
3.3	GPS SV visibility for a single angenna.	53
3.4	Visualization of GPS visibility on HEO.	54
3.5	Impact of our tropospheric model on SV visibility.	54
3.6	GPS Antenna Gain with Sidelobes	55
3.7	The Canadian Space Agency ground station in Saskatoon	56
4.1	Apollo 10 Lunar Module Rendezvous	65
5.1	Baseline simulation, position error, along/across/radial frame*.	86
5.2	Baseline simulation, velocity error, along/across/radial frame*.	86
5.3	Baseline simulation, receiver clock bias error and receiver clock mode.	87
5.4	Baseline simulation, receiver clock drift error and receiver clock mode.	87
5.5	Baseline simulation, position error and locked GPS satellites.	88
5.6	Baseline simulation, measurement residuals.	88
5.7	Zenith-pointing antenna, position error and locked GPS satellites.	92
5.8	Nadir-pointing antenna, position error and locked GPS satellites.	92
5.9	Simple propagator, position error.	94
5.10	Complex propagator, position error.	94
5.11	Sidelobes OFF, position error and locked GPS satellites.	96

5.12 Sidelobes ON, position error and locked GPS satellites.	96
5.13 Receiver Clock Perfect, position error.	98
5.14 Receiver Clock Perfect, position error and locked GPS satellites. . . .	98
5.15 Navigation using passive ranging and GPS, receiver clock perfect. . .	100
5.16 Navigation using passive ranging and GPS, receiver clock error included, $R_{rge} = (5 \text{ m})^2$	101
5.17 Navigation using passive ranging and GPS, receiver clock error included, $R_{rge} = (50 \text{ m})^2$	101
5.18 Unscented Kalman Filter, position error and locked GPS satellites . .	103
5.19 Continuous-Discrete Unscented Kalman Filter, position error and locked GPS satellites	103
A.1 Time Dilution of Precision over a single Molniya orbit	127

Nomenclature

List of Acronyms

AGI	Analytical Graphics Inc.
ARC	Ames Research Centre
CDCKF	Continuous-Discrete Cubature Kalman Filter
CDUKF	Continuous-Discrete Unscented Kalman Filter
CKF	Cubature Kalman Filter
DE	Developmental Ephemeris
DLR	German Aerospace Centre
DORIS	Doppler Orbitography and Radiopositioning Integrated by Satellite
ECEF	Earth-Centred Earth-Fixed
ECI	Earth-Centred Inertial
EGM-96	Earth Gravitational Model 1996
EKF	Extended Kalman Filter
FK5	5th Fundamental Catalog
GDGPS	Global Differential GPS
GDOP	Geometric Dilution of Precision
GEO	Geostationary Orbit
GEONS	GPS-Enhanced Onboard Navigation System

GNC	Guidance, Navigation, and Control
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HEO	Highly Elliptical Orbit
HPOP	High Precision Orbit Propagator
IMU	Inertial Measurement Unit
ITRF93	International Terrestrial Reference Frame 1993
JPL	Jet Propulsion Laboratory
LEO	Low Earth Orbit
LORAN	Long Range Navigation
MEO	Medium Earth Orbit
PCW	Polar Communication and Weather
PRN	Pseudo-Random Noise
RMS	Root Mean Square
SRP	Solar Radiation Pressure
STDP	Space Technology Development Program
STK	Satellite Tool Kit
SV	Space Vehicle
TAP	Triple Apogee (orbit)
TDOP	Time Dilution of Precision
TOB	Truth Orbit
UKF	Unscented Kalman Filter
UTC	Coordinated Universal Time
WGS-84	World Geodetic System 1984

List of Symbols

Note: Throughout this thesis, scalar values are typeset in italics (t , U , σ), vectors are lowercase boldface (\mathbf{x} , \mathbf{y} , \mathbf{s}), and matrices are uppercase boldface (\mathbf{P} , \mathbf{S} , $\mathbf{\Gamma}$).

Introduction

- \mathbf{f} physical model
- \mathbf{h} measurement model
- \mathbf{r} spacecraft position (ECI)
- t time
- \mathbf{v} spacecraft velocity (ECI)
- \mathbf{x} spacecraft state vector
- \mathbf{y} measurement vector
- σ standard deviation

Propagator Variables

- A_D atmospheric drag surface area
- A_{SR} solar radiation pressure surface area
- C_D atmospheric drag coefficient
- C_{SR} solar radiation pressure coefficient
- R_\oplus radius of the Earth, typical value 6378 km
- U gravitational potential function
- \mathbf{a} spacecraft acceleration (ECI)
- \mathbf{a}_{2body} two-body acceleration
- \mathbf{a}_3 third body acceleration
- \mathbf{a}_{drag} atmospheric drag acceleration

\mathbf{a}_{geoid}	geoidal Earth gravitational acceleration
\mathbf{a}_{SR}	solar radiation pressure acceleration
h	altitude
m	mass of satellite
p_{SR}	solar radiation pressure
\mathbf{r}	spacecraft position (ECI)
$\mathbf{r}_{X \rightarrow Y}$	position of Y with respect to X
t	time
\mathbf{v}	spacecraft velocity (ECI)
\mathbf{v}_{rel}	velocity of satellite relative to atmosphere
μ	gravitational parameter
μ_{\oplus}	gravitational parameter of the Earth, typical value $398\,600\,\text{km}^3\cdot\text{s}^{-2}$
ρ	atmospheric density
τ	time (used as the argument of integration)
ω_{\oplus}	angular velocity of Earth, typical value $7.29 \times 10^{-5}\,\text{rad}\cdot\text{s}^{-1}$
\odot	Sun
\oplus	Earth
\mathbb{C}	Moon

Measurement Variables

C/N_0	link budget carrier to noise ratio
G_{tx}	link budget transmitter gain
G_{rx}	link budget receiver gain
IF	ionosphere factor
L_{fs}	link budget free space loss

L_t	link budget temperature loss
P_{tx}	link budget transmitter power
\mathbf{Q}_{drift}	drifting receiver clock process noise
b	receiver clock bias
c	speed of light, $299\,792.458\,\text{km}\cdot\text{s}^{-1}$
d	receiver clock drift
\mathbf{h}	measurement model
\mathbf{h}_{GPS}	GPS pseudorange measurement
\mathbf{h}_{GPS}^Δ	GPS pseudorange measurement with ephemeris errors
\mathbf{h}_{iono}^f	ionospheric filter model
\mathbf{h}_{iono}^t	ionospheric truth model
\mathbf{h}_{Rclk}	receiver clock error model
q_1	drifting receiver clock bias noise
q_2	drifting receiver clock drift noise
\mathbf{r}	spacecraft position (ECI)
$\hat{\mathbf{r}}$	estimated spacecraft position
$\mathbf{r}_{SV(i)}$	GPS space vehicle position (ECI)
$\mathbf{r}_{SV(i)}^\Delta$	GPS SV position (ECI) with ephemeris errors
t	time
t_b	time of GPS signal broadcast
t_{GPS}	GPS system time
t_{Rclk}	receiver clock time
$t_{SV(i)}$	GPS SV clock time
w_b	drifting receiver clock bias noise

w_d	drifting receiver clock drift noise
w_s	steered receiver clock bias noise
\mathbf{x}	state vector
$\hat{\mathbf{x}}$	estimated spacecraft state
\mathbf{y}	measurement vector
$\hat{\mathbf{y}}$	estimated measurement vector
$\Delta\mathbf{y}_{iono}$	ionosphere error measurement vector
$\Delta\mathbf{y}_{Rclk}$	receiver clock error measurement vector
$\Delta\mathbf{y}_{SVclk}$	SV clock error measurement vector
σ_s	steered receiver clock standard deviation
τ_s	steered receiver clock decay constant

Filter Variables

Note: For filter variables, the subscript k indicates a discrete-time variable at time t_k and the superscript $-/+$ indicates an *a priori* or *a posteriori* variable, respectively.

$\mathbf{F}(\mathbf{x}, t)$	linearized physical model
\mathcal{F}	propagated state sigma points
$\mathbf{H}(\mathbf{x}, t)$	linearized measurement model
\mathbf{H}_k	measurement model matrix
\mathbf{I}	identity matrix
\mathbf{K}	Kalman gain
\mathbf{P}	state covariance matrix
\mathbf{P}_{xy}	state-measurement variance
\mathbf{P}_{yy}	measurement covariance
\mathbf{Q}	process noise covariance

R	measurement noise covariance
S	Cholesky decomposition of state covariance
\mathcal{X}	state vector sigma points
\mathcal{Y}	measurement sigma points
f	physical model
$\hat{\mathbf{f}}$	continuous-discrete state model
h	measurement model
\mathbf{s}_i	<i>i</i> th column of S
v	measurement model error
w	physical model error
w_i^c	sigma point weights, covariance
w_i^m	sigma point weights, mean
x	state vector
$\hat{\mathbf{x}}$	state estimate
y	measurement vector
$\hat{\mathbf{y}}$	estimated measurement
Γ	continuous-discrete covariance model
$\Phi(\mathbf{A})$	continuous-discrete matrix function
Φ_k	state transition matrix
α	UKF parameter
β	UKF parameter
κ	UKF parameter
λ	UKF parameter
τ	time (used as the argument of integration)

Chapter 1

Introduction

1.1 Motivation for this study: Satellite Service in Canada's North

The second-largest country in the world, Canada dominates the northern half of North America. However, much of the Canadian landscape is not arable, limiting significant settlement to a strip along the country's southern border [1]. The three northern territories, which together comprise a third of Canada's land mass, contain a mere 0.3% of the country's population [2]. These 100,000 Canadians live in isolated settlements that dot the North from the 60th parallel all the way to the tip of Ellesmere Island – at 83.5°N, the outpost of Alert, Nunavut is the northernmost permanently inhabited settlement in the world [3].

A G8 nation with a developed economy, Canada has undertaken to provide all of its citizens with basic services. Among these services are the means to communicate with the world around them and timely access to accurate weather information. Geostationary satellites are well-suited to provide these services, particularly in sparsely populated areas. However, satellites in geostationary orbit (GEO) have poor visibility of high latitudes due to their equatorial orbit, with signals essentially unavailable beyond 80° latitude (see Figure 1.1). Thus, communication service for the North is currently achieved using large, static ground stations and a terrestrial system of repeaters – Alert, for example, is serviced by a ground station at Eureka, Nunavut [4] – while weather services are carried out by American satellites flying in Low Earth Orbit (LEO) with a data latency up to six hours [5] (see Figure 1.2).

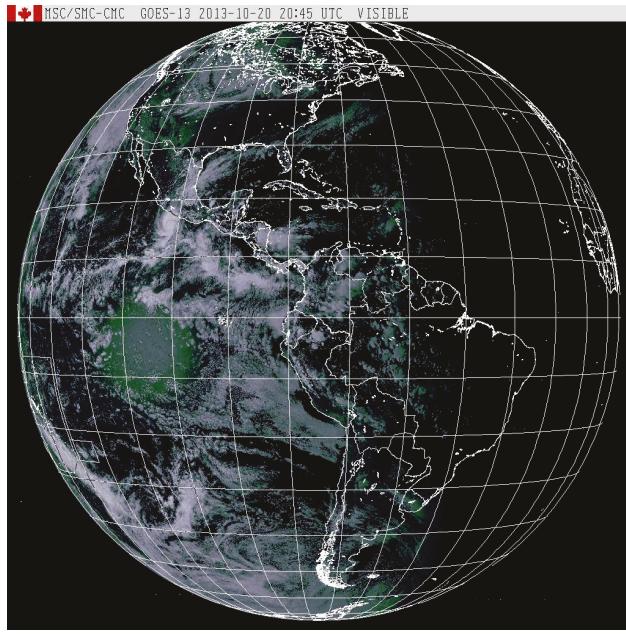


Figure 1.1: A sample image from GOES-13, an American weather satellite in geo-stationary orbit. Note the poor visibility of polar regions from this orbit.

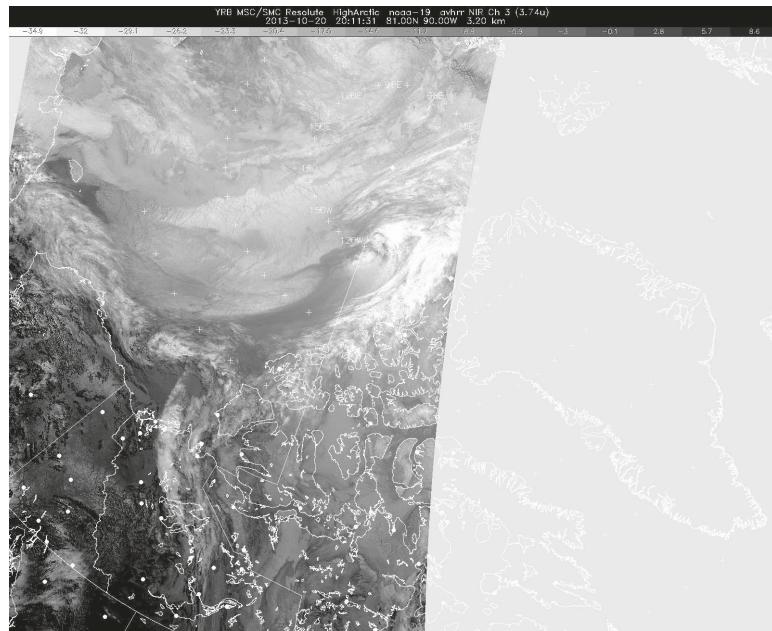


Figure 1.2: Imagery obtained from one pass of NOAA-19, an American weather satellite in low earth orbit. Note the incomplete ground coverage of this pass.

(Public domain imagery courtesy of Environment Canada and the U.S. National Oceanic and Atmospheric Administration: http://weather.gc.ca/satellite/index_e.html, 20 Oct 2013.)

1.2 Polar Communication and Weather

In recent years, the Canadian government has renewed its interest in Canada’s North as a means for economic development. In his 2013 Speech from the Throne, Governor General David Johnston affirmed the government’s commitment to: “securing our Northern sovereignty; promoting prosperity for Northerners; [and] protecting our Arctic environmental heritage.” [6].

In response to these past and present policy decisions, the Canadian Space Agency has proposed a new satellite mission. This mission, titled *Polar Communication and Weather (PCW)*, was first defined in 2007 and calls for a two satellite constellation. These satellites would be placed in a Molniya-type orbit (see Section 1.2.1) to maximize their ability to provide communication and weather services to Canada’s North. Once both satellites are in orbit, they would provide continuous communication coverage to Canada’s Arctic north of 70°N [7] and meteorological data on a 15 minute refresh rate for the entire circumpolar region north of 50°N [8]. The satellites are also specified to carry space weather instrumentation; these efforts would improve Canadian responsiveness to heightened geomagnetic activity and its effect on satellite operation and high-frequency communication in the Arctic [7].

In addition to its support of economic activity and environmental monitoring, PCW would positively affirm Canadian sovereignty in the North. In an age of observed climate change, navigation of the Northwest Passage (the historically proposed route from Europe to Asia through the Canadian Arctic) is becoming more feasible as the Arctic ice cap continues to recede [9]. PCW would permit active monitoring of the North and demonstrate (through public and private investment) the Canadian commitment to development in the North. In addition, its communication payload would be suitable for mobile users such as “ships, planes, and unmanned aerial vehicles” [7] – greatly assisting future Canadian Forces operations in the Canadian Arctic.

1.2.1 The Molniya Orbit

The *Molniya* orbit (from Russian молния, “lightning”) was conceived by the USSR for communication and reconnaissance during the Cold War. It is a highly-inclined, highly-elliptical orbit with a period of approximately twelve hours. The high inclination gives it access to high latitudes; the high ellipticity allows maximal coverage of the sub-satellite point at apogee; and the twelve-hour period gives the satellite a

repeating ground track with two main focus points separated by 180° in longitude. Thus, a single reconnaissance satellite with its perigee in the southern hemisphere and a controlled ascending node can cover the Soviet Union and United States every 24 hours, while a small constellation of communication satellites could offer northern Russia near-complete coverage when supported by appropriate ground hardware. Illustrations of this orbit can be found in Figure 1.3 (orbit in inertial space) and Figure 1.4 (orbit ground track).

The Molniya orbit is nominally inclined at the *critical inclination* of $i = 63.4^\circ$. It can be demonstrated that over time, the oblateness of the earth will cause the argument of perigee (see Appendix A.1) to rotate in the orbital plane (cf. de Ruiter et al. [10], pp. 168–169). However, the rate of this rotation is inclination-dependent. The critical inclination is the inclination that causes this perturbation to vanish, a desirable result for an orbit where perigee stability is essential to maintain coverage in the Northern hemisphere.

The Molniya orbit is just one example of a class of orbits known as Highly Elliptical Orbits (HEOs). Related high-inclination HEOs include:

- A twenty-four hour orbital period gives the **Tundra orbit** a repeating ground track with a single focus point (the sub-satellite point at apogee). At present, the only user of this orbit is Sirius Satellite Radio, which uses a trio of these satellites to maximise coverage of their market in North America [11].
- A sixteen hour orbital period gives the **Triple Apogee (TAP) orbit** a repeating ground track with three focus points separated by 120° in longitude. It has been proposed for PCW as a favourable alternative to the Molniya orbit as its higher perigee altitude lessens radiation exposure [12]. However, that discussion is beyond the scope of the present study.

In this study, we will focus on the Molniya orbit.

1.3 The Need for Navigation

Having established a need (communication and weather service for Canada’s North) and a means (a multi-purpose Canadian satellite on Molniya orbit), we now turn to

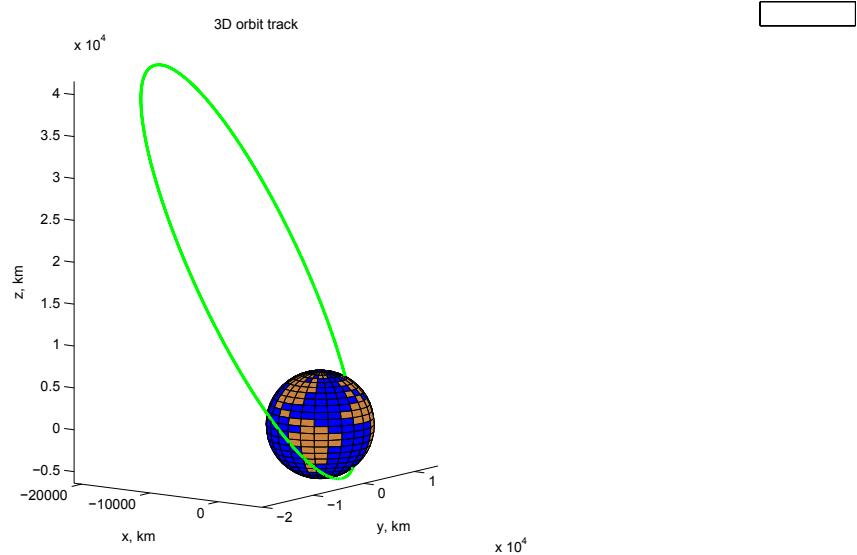


Figure 1.3: A typical Molniya orbit.

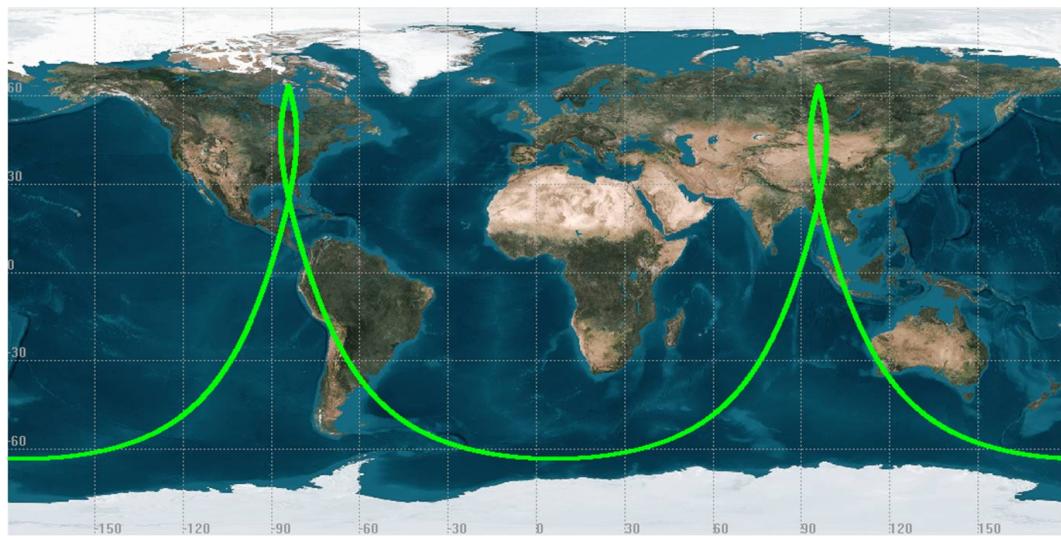


Figure 1.4: A typical Molniya ground track. Image generated using AGI STK 9.2.0.

one detail of implementation: spacecraft navigation. For our purposes, navigation is defined as:

*“The measurement and computation necessary to determine
the present spacecraft position and velocity.”*

– Apollo GNC Progress Report [13]

Knowledge of position and velocity is essential for any operational spacecraft. It is used to schedule uplink and downlink, a non-trivial conversation when one of the two parties is moving at multiple kilometres per second. It is used to plan, execute, and review spacecraft operations, both active (attitude manoeuvres, orbit correction, momentum dumping) and passive (imaging, in-situ measurements, battery charging). Last – but certainly not least – it is used to ensure spacecraft safety through advance warning of conjunction with orbital debris or other satellites.

Beyond this general need for navigation , the two primary roles of PCW – namely, communications and weather – are fundamentally dependent on knowledge of the spacecraft position at all times. Though the Molniya orbit permits a slow-moving sub-satellite point at apogee, precise ground-station pointing will be vital for reliable communication over some 40 000 km of slant range. This is particularly true for mobile users (see Section 1.2) with low-power ground stations. Similarly, the meteorological payloads for PCW will require accurate position knowledge as the spacecraft will be executing position-based attitude adjustments throughout the orbit to ensure an accurate meteorological field-of-view. We can therefore say with certainty that a navigation solution is critical to the future success of a PCW-like mission.

1.3.1 The Next Step: Autonomous Navigation

In recent years, spacecraft engineers have striven to design spacecraft capable of operation at varying levels of *autonomy* – that is, on-orbit operation without the intervention of a ground station. Autonomous navigation solutions offer distinct benefits for a spacecraft and its operators. These include:

- immediate availability of position and velocity estimates based on present measurements, rather than estimates based on previous ground passes.
- simplifying uplinks by reducing the need for ground stations to update the spacecraft with navigation information.

- increased operational flexibility by giving the spacecraft the ability to continue operations through a ground station outage.
- increased efficiency of the spacecraft by allowing it to prioritize tasks based on current and future orbital conditions.
- reduced overhead at the ground station, ultimately saving costs for the satellite operator.

In light of these benefits, autonomous navigation has been attempted and achieved by a number of spacecraft on low Earth orbit. Many of these spacecraft relied on signals from the Global Positioning System (GPS), including PoSAT-1 in 1993 (the first to autonomously generate its own orbital elements using GPS [14]) and CanX-2 in 2008 (the first CubeSat to successfully deliver GPS navigation fixes [15]).

1.3.2 Autonomous Navigation in HEO

Following these successes, the Canadian Space Agency expressed interest in extending these LEO-proven technologies to PCW on its highly elliptical orbit. However, navigation for HEO satellites has been not yet been achieved autonomously, with mission architects preferring the traditional method of ground-based ranging and tracking. Examples include the Sirius Satellite Radio constellation, which uses ground stations for tracking [16], and the Radiation Belt Storm Probes, which have used S-band Doppler based navigation since their launch in August 2012 [17]. This reliance on ranging is not without merit; when incorporated into a filter (see Section 1.5.3), such techniques can achieve accuracies on the order of 10 m (1σ) [18], more than accurate enough for typical communication and earth-observation missions.

While no HEO mission has yet performed autonomous navigation, past missions have demonstrated key capabilities:

- In 1997, a pair of experimental HEO satellites (EQUATOR-S and TEAMSAT) successfully received GPS signals at altitudes above that of the GPS constellation itself [19]. However, the reception was not autonomous as the process of signal acquisition was commanded from the ground.
- A classified American satellite launched in the 1990s used GPS signals to obtain a navigation solution in GEO [20]. The solution was achieved using ground-based processing in near-real time.

- The AMSAT-OSCAR 40 spacecraft (launched in 2000) autonomously acquired GPS signals throughout its 19-hour HEO [19]. Point solutions accurate to 3 km were achieved using ground-based processing.
- In early 2013, the GIOVE-A navigation mission became the first civilian satellite to achieve GPS navigation (position, velocity, and time) in a high altitude orbit [21]. This was achieved using sidelobe signals¹ as the satellite is in a circular orbit 1000 km above the GPS constellation.

Bolstered by this progress, the Canadian Space Agency specified autonomous navigation as a requirement for the PCW mission. This provided an opportunity for the growth of Canadian expertise in autonomous satellite navigation.

1.3.3 The Magellan Study on HEO Navigation

In 2011, the Canadian Space Agency contracted Magellan Aerospace of Ottawa to complete a Space Technology Development Program (STDP) from 2011–2014. This navigation-focused thesis was funded in part through a collaboration between Magellan Aerospace and Carleton University. Throughout this document, reference will be made to parameters of simulation that were defined to match those used by Magellan in their study. This parameter-matching allowed for direct comparison of simulation results as both studies attempted to meet the STDP navigation accuracy requirement of ± 150 m (3σ) in the along/across/radial reference frame². However, reference will not be made to the results of the Magellan study due to issues of confidentiality.

1.4 Problem Definition: Navigation on HEO

We are now at a point where we can define the motivating question for this study:

To what accuracy can we autonomously navigate a spacecraft in HEO?

For the sake of this study, we will adopt the operating parameters of the previously discussed Polar Communication and Weather mission (see Section 1.2), including its specified navigation requirements (see Section 1.3.3). Thus, we will examine a satellite

¹For a further discussion of GPS sidelobes, see Section 3.6.3.

²For a discussion of this reference frame, see Note in Section 5.3.1.

orbiting in a Molniya-type orbit and evaluate its ability to autonomously determine its position and velocity. For such an evaluation, we will require a *truth*.

Throughout this study, we will refer to this idea of truth, most notably in the form of a truth orbit. This truth orbit can be obtained in two ways. First, it can be generated using actual measurements obtained by a satellite in highly elliptical orbit. These measurements are post-processed on the ground, producing a truth orbit accurate to the limit of the measurement errors. Second, a trusted high-fidelity orbit propagator is employed to generate a precise truth orbit. We can then corrupt this truth with known errors, allowing great insight into the behaviour of our navigation solution for different error types. This level of control is entirely appropriate to the goals of the current study and will be our method of choice as we evaluate the sensitivity and accuracy of our proposed solution. For more details on the software used in this process, please see Appendix C.

1.5 Three Keys to Navigation

The answer to the motivating question in Section 1.4 – i.e., the accuracy to which we can navigate our spacecraft – will intrinsically rely on the following three components:

1. prediction of future spacecraft position and velocity based on our current knowledge of the spacecraft;
2. selection and modelling of measurements available to our spacecraft which contain information about the current spacecraft position and velocity; and
3. fusion of the prediction and the measurement information to achieve an accurate navigation solution.

These three topics will form the backbone of this study. They are introduced below and discussed in greater detail in their respective chapters (Chapters 2, 3, 4).

Before we discuss these three topics, we will introduce the concept of a *state vector*. The state vector is a mathematical construct which contains within it all information necessary to specify the current state of a given dynamic system. In this study, we will present a dynamic model appropriate for an Earth-orbiting spacecraft. In this dynamic model, we will model our satellite as a particle with its mass centred at the

spacecraft centre-of-mass. Our state vector, denoted $\mathbf{x}(t)$, can therefore be composed of the satellite's position and velocity:

$$\mathbf{x}(t) = [\mathbf{r}(t)^T \ \mathbf{v}(t)^T]^T = [r_x(t) \ r_y(t) \ r_z(t) \ v_x(t) \ v_y(t) \ v_z(t)]^T \quad (1.1)$$

where position $\mathbf{r}(t)$ and velocity $\mathbf{v}(t)$, both functions of time t , are typically broken down into their Cartesian components.

During this study, we may choose to append additional elements to our state vector. The states listed above, however, are the minimum required to characterize the dynamics of a particle and will be included in all aspects of our analysis.

1.5.1 Predicting: refining our expectations

The ability to predict the motions of celestial bodies has been sought for centuries. Ancient astronomers developed complex models of cycles and epicycles to reconcile the erratic motion of planets (from Greek $\pi\lambda\alpha\nu\eta\tau\eta\varsigma$, “wanderers”) with the Ptolemaic model of an Earth-centred universe. These models were standard until Kepler presented his three laws of orbital motion, laws which had theoretical grounding in Copernicus' heliocentric solar system model and practical grounding in Brahe's detailed astronomical observations. These laws were later refined by Newton to produce a mathematical description of gravity, the force responsible for keeping celestial bodies on their orbits.

If we consider two particles of mass m_1 and m_2 , the mutual gravitational force between them has a magnitude given by:

$$F_{grav} = G \frac{m_1 m_2}{|\mathbf{r}_{1 \rightarrow 2}|^2} \quad (1.2)$$

where $\mathbf{r}_{1 \rightarrow 2}$ is the position vector between the two particles and G is a constant of proportionality (Newton's gravitational constant, $6.67 \times 10^{-11} \text{ N}\cdot\text{m}^2\cdot\text{kg}^{-2}$).

Though knowledge of orbital dynamics has greatly developed since the time of Newton, the art of orbit prediction – also known as orbit propagation – remains reliant on this Newtonian ideal. To propagate an orbit, we attempt to solve the differential equation:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \quad (1.3)$$

with initial condition $\mathbf{x}(t_0)$, where \mathbf{x} is our previously defined state vector. The

function $\mathbf{f}(\mathbf{x}, t)$ is dominated by Newtonian gravity, yet includes within it the effects of a variety of perturbative forces.

We can therefore solve (1.3) to determine the time history of $\mathbf{x}(t)$. This can be equivalently described as orbit propagation (i.e. moving our satellite forward on its orbit) or a time update (calculating the state $\mathbf{x}(t_1)$ at some future time $t_1 > t_0$). The latter interpretation will prove essential to our navigation exercise. For further details, including the nature of the perturbative forces and mathematical treatment of all forces involved, see Chapter 2.

1.5.2 Measuring: selecting our observations

Our understanding of the natural world is reliant on our ability to make observations and understand how those observations relate to the system they describe. Whether we consider sextant-based navigation, infrared range measurements between docking spacecraft, or any of a host of scenarios, the measurements available are direct functions of the system state. The relationship between states and measurements can be formalized as:

$$\mathbf{y} = \mathbf{h}(\mathbf{x}, t) \quad (1.4)$$

where $\mathbf{h}(\mathbf{x}, t)$ is some function which calculates measurements \mathbf{y} based on the current state \mathbf{x} and time t .

While (1.4) is useful for our understanding of the system under study, we are typically interested in the inverse problem. Namely, what does a measurement $\mathbf{y}(t_i)$ at time t_i tell us about the state $\mathbf{x}(t_i)$?

The present study is therefore faced with a pair of tasks: determining which measurements will be available to our autonomous navigation solution, and determining what information they hold about the state of our satellite at the time of the measurement. These tasks – and the development of appropriate mathematical models – are expanded upon in Chapter 3.

1.5.3 Filtering: fusing for navigation

It has become apparent that we will have two separate means by which to estimate our spacecraft state: propagating and measuring. Both processes are subject to their own errors and inaccuracies. The question then arises: how do we reconcile an erroneous prediction of our spacecraft state with an erroneous measurement of spacecraft state?

Furthermore, is there any means by which we can incorporate previous predictions and measurements?

The answer to both questions lies in the art of filtering. Filtering is the fusion of predictions and measurements – of time updates and measurement updates – in the quest for an accurate navigation solution. When properly formulated, filters allow the estimation of state based on initial conditions, a dynamic model, a measurement model, and measurements - all of which have associated errors.

The classical sequential filtering algorithm for dynamic systems is the Kalman filter, first presented by Rudolf Kalman in his seminal 1960 article [22]. A detailed discussion of his algorithm and its derivatives is left to Chapter 4 – at present, we will state only that the intrinsic non-linearity of satellite navigation gives a great deal of opportunity for innovative methods of filtering.

1.6 Organization of this thesis

This thesis has five chapters beyond the current Introduction. In Chapter 2, we define the physical models used in this study, detailing each of the physical forces which act on a spacecraft in highly elliptical orbit. In Chapter 3, we define the measurement models used in this study, including a thorough discussion of GPS signals and their use for navigation in HEO. In Chapter 4, we summarize the Kalman filter and discuss various ways by which it may be applied to nonlinear systems. In Chapter 5, we incorporate techniques from the three previous chapters to realize a HEO-appropriate navigation algorithm; the performance of this algorithm is then examined through simulation. Finally, in Chapter 6 we draw conclusions on the performance of our algorithm and provide recommendations for future autonomous navigation in highly elliptical orbit.

Chapter 2

Physical Models and Propagation

2.1 Introduction

2.1.1 Dead Reckoning: The Art of Propagation

Fundamentally, the art of propagation is a process of predicting the future based on knowledge of the past. In our current study, we are concerned with the prediction of spacecraft position and velocity over time. Historically, this prediction has been achieved using an orbit propagator, an algorithm designed to predict the positions of celestial bodies.

The propagator operates on the principle of “dead reckoning”: prediction of the future using information collected at present. This process traces its roots back to the age of sail, where sailors would measure the rate at which a rope would pay out when a log tied to one end of the rope was thrown overboard (see Fig. 2.1). The chip log reading, a measurement of ship speed, was combined with knowledge of the local environment (e.g., prevailing winds) to produce an estimate of forward travel until the next reading.

In this ship-borne example, velocity is estimated to determine a change in position by simple integration. Orbit propagators go one step further, modelling the acceleration acting on a spacecraft to determine changes in velocity, then using this knowledge to determine changes in position. From kinematics, we recall the following relationships:

$$\mathbf{v}(t) = \mathbf{v}_0 + \int_{t_0}^t \mathbf{a}(\tau) d\tau \quad (2.1a)$$

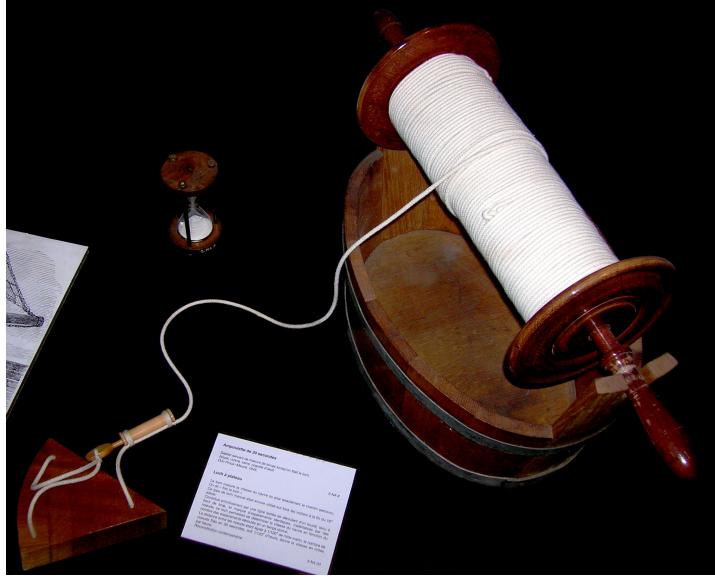


Figure 2.1: A “chip log”, used for shipborne speed estimation.
©Rémi Kaupp [23]

$$\mathbf{r}(t) = \mathbf{r}_0 + \int_{t_0}^t \mathbf{v}(\tau) d\tau \quad (2.1b)$$

To solve this initial value problem for the time histories of velocity $\mathbf{v}(t)$ and position $\mathbf{r}(t)$, we need two things:

1. the initial conditions $\mathbf{r}_0 = \mathbf{r}(t_0)$ and $\mathbf{v}_0 = \mathbf{v}(t_0)$;
2. the acceleration $\mathbf{a}(t)$ acting on our spacecraft for all times t .

In the case of a spacecraft, the initial conditions are determined using measurements. These measurements can be ground-based (e.g. optical observations) or space-based (e.g. reception and processing of signals from the Global Positioning System). This process, known as *initial orbit determination*, has a significant impact on the accuracy of predictions $\mathbf{v}(t)$ and $\mathbf{r}(t)$. However, initial orbit determination is beyond the scope of the present study.

Instead, this chapter will focus on the latter requirement, the development of acceleration models $\mathbf{a}(t)$ specific to the spacecraft environment. As we develop these models, it will become apparent that for spacecraft, the time-dependent accelerations will also depend on spacecraft state $\mathbf{x}(t) = [\mathbf{r}(t) \ \mathbf{v}(t)]^T$. The task at hand is therefore the determination of $\mathbf{a}(t, \mathbf{x})$, the dependence of spacecraft acceleration on its current position, its current velocity, and the time-dependent orbital environment.

Note on form of the propagator: The solution of (2.1) implies the use of some method of integration. We recall from mathematics that integrals can be evaluated both analytically and numerically. Orbit propagators can thus be divided into two broad categories: analytical and numerical. As discussed by Montenbruck [24], standard analytical propagators such as the SGP4 model are limited to kilometre-level accuracy due to the neglect of short period effects. Since our current study is concerned with short periods (i.e., less than one orbit), numerical integration of the equations of motion is in order. For details on the form of the numerical integrator used in this study, please consult Appendix C.2.

2.1.2 Accelerations at play on the HEO orbit

As discussed above, our ability to predict the future state of our spacecraft is dependent on our knowledge of the acceleration $\mathbf{a}(t, \mathbf{x})$ acting on the spacecraft. In the idealized analysis, the motion of Earth-orbiting satellites is modelled by the two-body problem. To solve this problem, we make two assumptions:

1. both the Earth and the satellite are point masses¹;
2. the central gravitational force of one body is the *only* force acting on the other body, and vice versa.

Having made these assumptions, we can use (1.2) (see Section 1.5.1) to derive the celebrated *two-body equation* (cf. Vallado [26], pp. 20–23). Further, since the mass of one body (the Earth) dwarfs the mass of the other body (the satellite), we can simplify the two-body acceleration to:

$$\mathbf{a}_{\text{2body}} = -\mu_{\oplus} \frac{\mathbf{r}}{r^3} \quad (2.2)$$

where \mathbf{r} is the position vector from the mass centre of the Earth to the mass centre of the orbiting body, and μ_{\oplus} is the standard gravitational parameter of the Earth.

Fortunately for humanity, our Earth is not a featureless and unchanging point mass. We have sunlight by day, moonlight by night, an atmosphere to breathe, and plenty of terrain to explore. These departures from simplicity have an unavoidable impact on the motions of Earth-orbiting satellites. More specifically, the motions of satellites are affected by:

¹By the *shell theorem*, this is equivalent to assuming a spherically symmetric mass distribution for the Earth (Newton [25], Book I, Section XII, Theorem XXXI).

- non-uniformity of the Earth’s gravitational field;
- the gravitational pull of the Sun, Moon, and planets;
- drag forces during travel through the Earth’s atmosphere;
- momentum transferred by sunlight incident on the spacecraft;

among others. Each of these effects can be modelled as an acceleration or force (see Note below) which “perturbs” the solution predicted by the two-body acceleration. The perturbations are added to the two-body acceleration in (2.2), with the resulting acceleration then used in (2.1) to perform the orbit propagation.

In the remainder of this chapter, we will discuss and present models for each perturbing acceleration. We will then implement the model in a MATLAB environment and verify this implementation against a known truth.

Note on “Forces” and “Accelerations”: We recall that Newton’s Second Law of Motion for a single body can be stated as $\Sigma \mathbf{F} = d\mathbf{p}/dt$, where \mathbf{F} are external forces acting on the body and \mathbf{p} is the momentum of the body. This simplifies to the more familiar $\Sigma \mathbf{F} = m\mathbf{a}$ for the case of unchanging mass m , where \mathbf{a} is the acceleration experienced by the body. Since we are considering forces acting on a simple satellite model (including an unchanging mass), discussions of “perturbing accelerations” and “perturbing forces” are equivalent.

2.2 Perturbing Accelerations and Their Models

In this section, we will discuss the four major accelerations (the *primary perturbations*) that act on Earth-orbiting satellites and present a mathematical model for each perturbation. We will also introduce a number of smaller perturbations (*secondary perturbations*) that have a measurable impact on a satellite in HEO.

2.2.1 Earth Gravity

Attempts to characterize the gravitational field of the Earth have been underway since the Scientific Revolution. As early as 1672, pendulum-based measurements of gravitational acceleration in France and French Guiana demonstrated that Earth’s gravitational field is non-uniform [27]. Measurements continued for hundreds of years, generating a wealth of location-specific gravity data.

These measurements were then used in the development of *geopotential models*, theory-based attempts to represent the gravitational field of the Earth as a whole. These models hinge on the nature of gravity as a *conservative* force. We recall from physics that such forces can be fully described by a scalar potential function. The force vector at a given point is then calculated by applying the gradient operator to the scalar potential at that point.

Spherical harmonic expansion of Earth gravitational potential

One such model, the spherical harmonic expansion, was first examined in 1937 by Dubovskii [28]. Spherical harmonics are a series of orthogonal functions defined for spherical coordinates. Just as Fourier theory allows us to construct any periodic function using a weighted summation of sines and cosines, we can construct any square-integrable spherical function using a weighted summation of spherical harmonics (cf. Groemer [29], p. 63).

Using the *associated Legendre polynomials*, a series of spherical harmonic functions detailed in the literature (e.g. Abramowitz and Stegun [30], pp. 331–341), we can construct a potential function $U(r, \phi_{gc}, \lambda)$ representing the Earth's gravitational potential. At present, the standard representation of U (see Vallado [26], p. 543) takes the form:

$$U(r, \phi_{gc}, \lambda) = \frac{\mu_{\oplus}}{r} \sum_{n=0}^{\infty} \sum_{m=0}^l \left(\frac{R_{\oplus}}{r} \right)^n \cdot P_{nm}[\sin(\phi_{gc})] \cdot \{C_{nm} \cos(m\lambda) + S_{nm} \sin(m\lambda)\} \quad (2.3)$$

where we define:

$U(r, \phi_{gc}, \lambda)$ gravitational potential function, $\text{km}^2 \cdot \text{s}^{-2}$

r radial coordinate, km

ϕ_{gc} geocentric latitude coordinate, rad

λ longitude coordinate, rad

μ_{\oplus} gravitational parameter of Earth, typical value $398\,600 \text{ km}^3 \cdot \text{s}^{-2}$

n summation index for model degree

m summation index for model order

R_\oplus	mean equatorial Earth radius, typical value 6378 km
P_{nm}	associated Legendre polynomials of order l and degree m
C_{nm}	gravity model coefficients, cosine terms ²
S_{nm}	gravity model coefficients, sine terms

To calculate the acceleration predicted by this potential function, we must take the gradient³:

$$\mathbf{a}_{\text{geoid}} = \nabla U \quad (2.4)$$

A *gravitational model* is a set of parameters $[\mu_\oplus \ R_\oplus \ C_{nm} \ S_{nm}]$ that defines U . In this study, we will use the *ellipsoid* $[\mu_\oplus \ R_\oplus]$ (see Table 2.7) defined by the World Geodetic System 1984 (WGS-84). This ellipsoid is the reference used by the Global Positioning System (GPS), a system on which the navigation solution in this study will rely. We will use the *geoid* coefficients $[C_{nm} \ S_{nm}]$ defined by the Earth Gravitational Model 1996 (EGM-96), which improve upon the geoid coefficients released with WGS-84 and are compatible with the older ellipsoid [31].

Once a gravitational model is selected, we define its fidelity by setting the summation indices n and m in (2.3). The maximum value for n is the *degree* of the model, while the maximum value for m is the *order* of the model. This degree and order are often written in the shorthand $n_{\max} \times m_{\max}$, e.g., a 10×10 EGM-96 model.

A recursive algorithm for Earth gravitational acceleration

While concise on paper, the harmonic series expansion stated in (2.3) is computationally intensive. Furthermore, the series itself only defines the gravitational potential; as stated in (2.4), the gradient of the potential must also be calculated, itself an involved process.

Fortunately, there exist recurrence relations for the associated Legendre polynomials P_{nm} . We can exploit these to define an efficient calculation routine for our

²The C_{20} term, which characterizes the Earth's oblateness, is the dominant term of the geoidal perturbation. Its negative is commonly known as the J_2 coefficient.

³Note that, by geophysical convention, the gravitational potential function is the negative of the gravitational potential energy. For this reason, we calculate the acceleration using the *positive* gradient, rather than the negative gradient as is common in physics. For further discussion, see Vallado [26], p. 520.

desired acceleration as a function of the current position. For details on the algorithm employed for this study, see Appendix B.

2.2.2 Third-Body Gravity

The phrase *third-body gravity* refers to the gravitational effect of a “third body” in the two-body system modelled by (2.2). For Earth-orbiting satellites, this third body is typically the Sun, the Moon, or one of the other planets in our Solar System.

Once again, we model each body as a point mass and neglect all forces save gravity. Using Newtonian mechanics, the third-body effect can be derived (cf. Berry and Coppola [32]) and expressed as:

$$\mathbf{a}_3 = \mu_3 \left(\frac{\mathbf{r}_{sat \rightarrow 3}}{r_{sat \rightarrow 3}^3} - \frac{\mathbf{r}_{\oplus \rightarrow 3}}{r_{\oplus \rightarrow 3}^3} \right) \quad (2.5)$$

where we define:

\mathbf{a}_3	third body acceleration on spacecraft, $\text{km}\cdot\text{s}^{-2}$
μ_3	gravitational constant of third body, $\text{km}^3\cdot\text{s}^{-2}$
$\mathbf{r}_{sat \rightarrow 3}$	relative position vector from satellite to third body, km
$\mathbf{r}_{\oplus \rightarrow 3}$	relative position vector from Earth to third body, km

Using this equation, we can examine the perturbing accelerations caused by each body in our solar system. The results of such an examination are shown in Table 2.1, which employs data from Vallado [26] (pp. 990–991).

It is apparent from Table 2.1 that the accelerations of the Moon and Sun dwarf those of the other planets. This is typical for Earth-orbiting satellites. It is therefore customary to only consider the third-body effects of the Moon and Sun. In this case, these effects are termed *luni-solar perturbations*. In our study, we include the luni-solar perturbations and will use the Jet Propulsion Laboratory Developmental Ephemeris (JPL DE-421 – see Appendix A.2) to obtain $\mathbf{r}_{\oplus \rightarrow \odot}(t)$, $\mathbf{r}_{\oplus \rightarrow \mathfrak{C}}(t)$, μ_\odot and $\mu_{\mathfrak{C}}$. (The latter can be found in Table 2.7.)

Table 2.1: Gravitational Influence of Celestial Bodies at Earth

	closest approach to Earth (AU)	gravitational parameter, $\text{km}^3\cdot\text{s}^{-2}$	worst-case acceleration, $\text{km}\cdot\text{s}^{-2}$
Sun	1.00	133 000 000 000	3.65×10^{-9}
Mercury	0.61	22 000	2.63×10^{-15}
Venus	0.28	325 000	4.21×10^{-13}
Moon	0.002 57	4 900	9.63×10^{-9}
Mars	0.52	42 800	8.19×10^{-15}
Jupiter	4.20	127 000 000	4.69×10^{-14}
Saturn	8.54	37 900 000	1.68×10^{-15}
Uranus	18.19	5 790 000	2.64×10^{-17}
Neptune	29.07	6 840 000	7.65×10^{-18}

Note: For the purposes of this table, we have defined “closest approach to Earth” as the difference between the semimajor axes of the Earth and the planet in question. The exceptions are the Sun (the semimajor axis of Earth is used) and the Moon (the lunar perigee is used). We then assume perfect alignment, which causes (2.5) to become scalar and allows us to set $r_{sat \rightarrow 3} = r_{\oplus \rightarrow 3} - r_{\oplus \rightarrow sat}$, where $r_{\oplus \rightarrow sat}$ is calculated at our HEO apogee (46 000 km).

2.2.3 Atmospheric Drag

One of the four fundamental forces of flight, drag resists the movement of an object through an enveloping fluid. Counter to our commonly-held notion of “the vacuum of outer space”, the Earth’s atmosphere is measurably present at altitudes typical for spacecraft near the Earth. As a first estimate, we can adapt the classical drag formula (cf. Vogel [33], p. 89):

$$\mathbf{a}_{drag} = -\frac{1}{2} \frac{C_D A_D}{m} \rho |\mathbf{v}_{rel}|^2 \hat{\mathbf{e}}_{v_{rel}} \quad (2.6)$$

where we define:

\mathbf{a}_{drag} drag acceleration on spacecraft, $\text{km}\cdot\text{s}^{-2}$

C_D drag coefficient, typical value 2.2 (Vallado [26], p. 549)

A_D surface area of satellite perpendicular to $\hat{\mathbf{e}}_{v_{rel}}$, km^2

m	mass of satellite, kg
ρ	local atmospheric density, kg·km ⁻³
\mathbf{v}_{rel}	velocity of satellite relative to atmosphere, km·s ⁻¹
$\hat{\mathbf{e}}_{v_{rel}}$	$\mathbf{v}_{rel}/ \mathbf{v}_{rel} $, unitless

Unlike the equation for Earth gravity, the drag equation appears straightforward to implement. However, this simplicity is deceptive. As outlined by Vallado [26] (pp. 549–553), the difficulty with satellite drag is determining the value of each parameter:

- Atmospheric density ρ decreases with altitude at a roughly exponential rate (cf. Wertz and Larson [34], p. 211). It can also vary by an order of magnitude depending on solar activity (*ibid.*, p. 984).
- Spacecraft mass m and surface area A_D can vary widely based on mission phase. The latter is also attitude dependent.
- Relative velocity \mathbf{v}_{rel} is, by definition, reliant on the velocity of the atmosphere – which in turn necessitates the need for a wind model. Current models are computationally intensive and only provide a rough estimate.
- Finally, the drag coefficient C_D is itself empirically determined and varies with satellite attitude and surface composition.

In response, we examined each of the issues discussed above.

Drag Simplified: atmospheric density ρ . Our current study is focused on an orbital regime (HEO) in which our spacecraft spends much of its time at high altitude. Since ρ falls off with altitude at a roughly exponential rate, drag makes a significant impact only at orbit perigee. For this reason, we have adopted a uniform layer model for the atmospheric density:

$$\rho(h) = \begin{cases} \rho_{500}, & \text{if } h \leq h_{\text{cutoff}} \\ 0, & \text{otherwise} \end{cases} \quad (2.7)$$

where altitude is defined as $h = r - R_{\oplus}$ ⁴, the density of this uniform layer is chosen to be the density at the nominal perigee altitude of 500 km (ρ_{500}), and the thickness of the uniform layer (h_{cutoff}) is reserved as a tuning parameter. A more complex exponential model for atmospheric density was also examined but was not found to provide increased propagator accuracy for this particular orbit.

In addition to its computational simplicity, this density model carries an additional benefit: it is not dependent on solar activity, removing the need for knowledge of the time-dependent solar radio flux and geomagnetic activity. As noted by Montenbruck et al. [35], this real-time data is rarely available to an onboard navigation scheme.

Drag Simplified: relative velocity. For the calculation of \mathbf{v}_{rel} we assume a “glued-on atmosphere” – i.e., one that co-rotates with the Earth. This allows us to avoid computationally intensive wind models and set $\mathbf{v}_{\text{rel}} = \mathbf{v} - \omega_{\oplus} \times \mathbf{r}$, where \mathbf{v} and \mathbf{r} describe the state of our spacecraft and ω_{\oplus} is the angular velocity of the rotating Earth.

Drag Simplified: satellite parameters. The evaluation of (2.6) is further simplified by holding satellite parameters A , m , and C_D constant. This implies averaging of the attitude-dependant parameters A and C_D over an orbit period.

The simplifications outlined above are appropriate to a real-time on-orbit implementation in line with the motivating case for this study. Furthermore, in simulation (see 2.3.2) this simplified model was found to be sufficiently accurate. This model is therefore implemented in the final propagator with values given in Table 2.7.

2.2.4 Solar Radiation Pressure

One measurable consequence of the particle nature of light is the existence of electromagnetic momentum. Quanta of light which strike any non-transparent surface impart an impulse (a change in momentum) onto that surface. This impulse is of such a small magnitude that it is insignificant for typical Earthbound kinematic analyses and is accounted for only in precise laboratory work.

⁴Note that this definition implicitly assumes a spherical model for the Earth. Given the other uncertainties in the drag model, this is an appropriate assumption.

However, in the nearly friction-free environment of Earth orbit, near-constant sunlight (from a near-constant direction) causes a very real change in momentum. Following the example of Vallado [26] (p. 577), we model this continuous change in momentum as a pressure upon the area of the satellite exposed to solar radiation:

$$\mathbf{a}_{SR} = -\frac{p_{SR}C_{SR}A_{SR}}{m} \hat{\mathbf{e}}_{r_{sat \rightarrow \odot}} \quad (2.8)$$

where we define

\mathbf{a}_{SR}	solar radiation acceleration on spacecraft, $\text{km}\cdot\text{s}^{-2}$
p_{SR}	solar radiation pressure, $\text{kN}\cdot\text{m}^{-2}$ ⁵
C_{SR}	surface interaction coefficient (see Table 2.2), unitless
A_{SR}	surface area of satellite perpendicular to $\hat{\mathbf{e}}_{r_{sat \rightarrow \odot}}$, m^2
m	mass of satellite, kg
$\mathbf{r}_{sat \rightarrow \odot}$	relative position vector from satellite to the Sun, km
$\hat{\mathbf{e}}_{r_{sat \rightarrow \odot}}$	$\mathbf{r}_{sat \rightarrow \odot}/ \mathbf{r}_{sat \rightarrow \odot} $, unitless

Like the drag acceleration (2.6), the solar radiation pressure acceleration is dependent on the mass and surface area of the satellite and features a surface-dependent coefficient (see Table 2.2). However, instead of opposing the satellite velocity, it always acts along the sun-satellite vector. Thus, while drag tends to lower perigee velocity (and thus reduce the semi-major axis of the orbit), solar radiation pressure can cause a change of orbital plane, among other effects.

Again, we face challenges determining the values of the parameters in (2.8). For the purposes of this study, we set C_{SR} , m and A_{SR} constant. These values can be found in Table 2.7.

2.2.5 Secondary Perturbations

The primary perturbations discussed above are just a few of the forces to which Earth-orbiting satellites are subjected. In this study, we also examined:

⁵Solar radiation pressure has a nominal value of $4.57 \times 10^{-9} \text{ kN}\cdot\text{m}^{-2}$ but varies with solar activity and with sun-satellite distance; see Appendix D for details.

Table 2.2: Radiation Pressure Coefficient

C_{SR}	Surface Type	Physical Interpretation
0.0	transparent	light passes through surface, no momentum transmitted
1.0	black body	light absorbed by surface, all momentum transmitted
2.0	ideal reflective body	light reflected back towards source, double momentum transmitted

- *solid Earth tides* – the redistribution of the Earth’s crust, mantle, and core effected by lunar and solar gravity;
- *ocean tides* – the analogous redistribution of the oceans;
- *Earth albedo pressure* – the extension of radiation pressure to include “Earth-shine”, sunlight reflected off the Earth;
- *Earth thermal pressure* – the extension of radiation pressure to include the infrared radiation of the Earth as a “grey body” with a non-zero temperature.

The tidal forces are modelled by modifying the $[C_{nm} S_{nm}]$ coefficients in (2.3) (cf. McCarthy and Petit [36]) while thermal and albedo pressure are modelled in a similar fashion to solar radiation pressure (cf. Klinkrad and Fritzsche [37]).

Anticipating that these forces would have a relatively minor impact on our navigation solution, we performed an analysis using Satellite Tool Kit (STK 9.2.0), an industry-standard software produced by Analytical Graphics Inc. (AGI). In this analysis, we set our initial conditions per Table 2.3 (corresponding to a 12-hour Molniya orbit). We then propagated a truth orbit⁶ by including the primary perturbations in accordance with Table 2.8. We then performed five further simulations: four including the primary perturbations and a single secondary perturbation in accordance with Table 2.9, and one including all primary and secondary perturbations. The period of analysis was 24 hours (two nominal Molniya orbit periods); for justification, see Note below. The results can be seen in Figure 2.2.

⁶A detailed definition of a truth orbit can be found in Section 2.3.

Table 2.3: Initial Conditions for Molniya Orbit. Orbital elements (see Appendix A.1) chosen to correspond to a Magellan Aerospace study [38].

semi-major axis	a	26 553.4 km
eccentricity	e	0.740969
inclination	i	63.4°
right ascension of ascending node	Ω	108.208°
argument of perigee	ω	270°
true anomaly	ν	0°
epoch	t_0	00:00:00 UTC, 04 Apr 2012

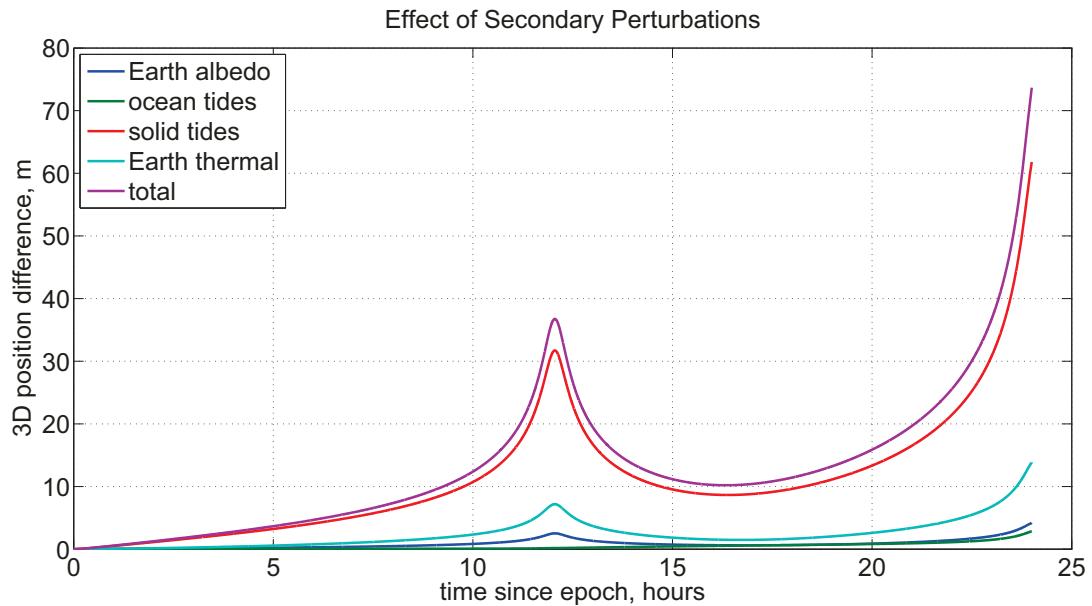


Figure 2.2: Position differences when secondary perturbations are modelled during propagation. Simulation begins at perigee and runs for two orbital periods; perigees thus occur at 0 h, 12 h, and 24 h.

In this figure, we see the 3D position difference between our truth orbit and the five modified orbits after twenty-four hours of propagation. This difference starts at zero (initial conditions are identical: the perigee of our Molniya orbit) and increases with each orbit. Maximal differences are observed at perigee (12 h, 24 h), a common result for Molniya orbits as this is the regime of maximum velocity.

It is readily apparent that solid tides cause the most significant modelling errors, resulting in approximately 62 m of position difference. When combined with ocean tides (3 m), albedo pressure (4 m), and thermal pressure (14 m), the effects combine for a total of 73 m of position difference. Bearing in mind the accuracy requirements of the current study (see Section 1.3.3) and the fact that our navigation solution will be updated by measurements over the orbit (especially at perigee), these secondary perturbations will not be considered for further analysis. This conclusion agrees with that of a similar analysis performed by Montenbruck et al. [35] for a satellite in LEO.

Note on the period of simulation: As demonstrated in Chapter 5 of this thesis, our navigation solution will never complete an entire orbit of propagation without some form of measurement-based correction. Thus, a two-orbit period for perturbation analysis is appropriate; it allows us to see the effects of each perturbation over a sufficiently long period of propagation (including a succession of apsis passes) while retaining the resolution to observe detailed perturbation effects over the entire orbit.

2.3 Verifying the Propagator

In the introduction to this chapter, we recalled the sailors of old as fellow navigators engaged in the art of propagation. Just as they used winds, currents, and charts to predict their path, so too have we developed models for our own orbital propagation. However, lest we sail for the Orient yet run aground in Hispaniola, it is vital to verify these models. For that, we must seek the truth.

In navigation studies, a *truth orbit* is an orbit generated by a trusted source to which a navigation solution can be compared. When analyzing data from a flown mission, the truth orbit is often a post-processed ground solution generated from accurate measurements such as satellite laser ranging. However, concept studies typically rely on trusted orbit propagators of high fidelity. Such propagators can produce

Table 2.4: Truth Orbits for Propagator Verification

Orbit ID	Included Forces	Corresponding Equation
TOB-1	$\mathbf{a} = \mathbf{a}_{2body}$	(2.2)
TOB-2	$\mathbf{a} = \mathbf{a}_{geoid}$	(2.3) (70×70), (2.4)
TOB-3	$\mathbf{a} = \mathbf{a}_{2body} + \mathbf{a}_{3,\odot} + \mathbf{a}_{3,\mathbb{C}}$	(2.5)
TOB-4	$\mathbf{a} = \mathbf{a}_{2body} + \mathbf{a}_{drag}$	(2.6)
TOB-5	$\mathbf{a} = \mathbf{a}_{2body} + \mathbf{a}_{SR}$	(2.8)
TOB-6	$\mathbf{a} = \mathbf{a}_{geoid} + \mathbf{a}_{3,\odot} + \mathbf{a}_{3,\mathbb{C}} + \mathbf{a}_{drag} + \mathbf{a}_{SR}$	

highly realistic orbits which are known to within machine precision, providing a precise standard to which we can compare our navigation solution. In addition, we have complete control over all error sources introduced, allowing us to individually examine the impact of each error on our solution. Thus, a simulated truth orbit is an appropriate tool for navigation solution validation.

Using STK’s High Precision Orbit Propagator (HPOP), we generated a number of truth orbits to validate our propagator. The six orbits, labelled **TOB-1** to **TOB-6**, are identified in Table 2.4. The forces included in the STK propagation are configured in accordance with Tables 2.7 and 2.8. Using these orbits, we will first examine the impact of each force (i.e., the position difference caused by each) to determine whether its model should be included in our navigation solution. We will then solve (2.1) using the models we developed above, examining the ability of our models to match the higher fidelity models in STK.

2.3.1 Impact of the Perturbing Accelerations

Before we examine each of the primary perturbations individually, it is worthwhile to determine just what impact they cause. For this purpose, we compared **TOB-2**, **-3**, **-4**, and **-5** to **TOB-1**. The results can be seen in Fig. 2.3.

In this figure, we notice that geoidal Earth gravity accounts for some 6500 km of 3D position difference after twenty-four hours of propagation, making it by far the most significant perturbation to our satellite. Beyond this, we observe that third-body gravity (23 km), drag (0.60 km), and solar radiation pressure (0.17 km) all cause

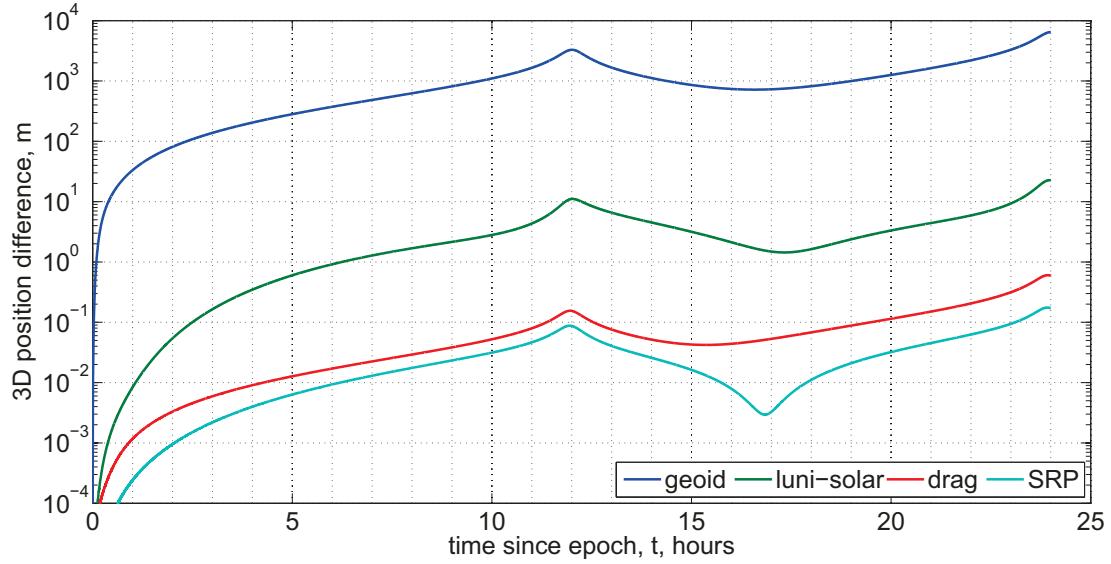


Figure 2.3: Position differences when primary perturbations are compared to the two-body solution. Simulation begins at perigee and runs for two orbital periods; perigees thus occur at 0 h, 12 h, and 24 h.

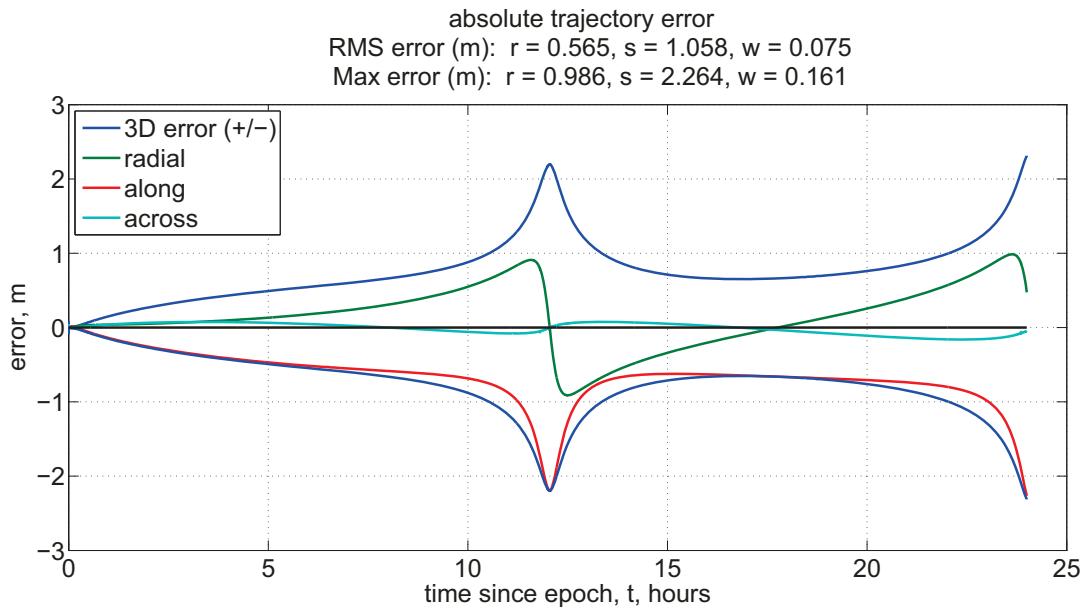


Figure 2.4: The performance of our propagator against the truth over two orbital periods. For details on the along/across/radial frame, see Note in Section 5.3.

Table 2.5: Propagator Verification Results.

3D position differences after twenty-four hours of propagation for:

- (i) perturbation truth orbits compared to two-body (**TOB-1**), and
- (ii) propagator models compared to corresponding truth orbits.

Perturbing acceleration	Truth orbit	(i) Perturbation impact, m	(ii) Model-Truth residual, m
\mathbf{a}_{geoid}	TOB-2	6 474 534	6.79
$\mathbf{a}_{3,\odot} + \mathbf{a}_{3,\mathbb{C}}$	TOB-3	22 653	1.28
\mathbf{a}_{drag} ($h_{cutoff} = 600$ km)	TOB-4	604	289.89
\mathbf{a}_{drag} ($h_{cutoff} = 865$ km)	TOB-4	604	7.00
\mathbf{a}_{SR}	TOB-5	174	2.28
all the above	TOB-6	6 453 150	2.31

differences in excess of our accuracy requirements (see Section 1.3.3). Thus, we have justified their inclusion in our propagator.

2.3.2 Propagator Accuracy

Now that we have developed models for the perturbations significant to our HEO orbit, we can measure them against the truth models described in Table 2.4. As our metric, we compared the 3D position difference after twenty-four hours (two nominal Molniya orbit periods) of propagation. Numerical results for **TOB-2** to **-6** can be found in Table 2.5, while **TOB-6** is visualized in Fig. 2.4.

Based on these simulations, some key results can be drawn.

1. Conservative forces (i.e. gravity) are modelled accurately;
2. As expected (cf. Section 2.2.3), drag was the most difficult force to model accurately. The model accuracy was improved substantially by using the reserved parameter (uniform atmosphere thickness h_{cutoff}) for tuning.
3. Solar radiation pressure is sufficiently accurate despite the lack of an Earth eclipse model.

Table 2.6: Summary of propagator models

Earth gravity	STK truth	EGM-96 70×70 on WGS-84 ellipsoid
	propagator	EGM-96 10×10 on WGS-84 ellipsoid $\mathbf{a}_{geoid} = \nabla U$ (2.3)
third-body gravity	STK truth	JPL DE-421 ephemerides, Sun and Moon
	propagator	JPL DE-421 ephemerides, Sun and Moon $\mathbf{a}_3 = \mu_3 (\mathbf{r}_{sat \rightarrow 3}/r_{sat \rightarrow 3}^3 - \mathbf{r}_{\oplus \rightarrow 3}/r_{\oplus \rightarrow 3}^3)$ (2.5)
atmospheric drag	STK truth	Jacchia-Roberts atmosphere
	propagator	uniform-layer atmosphere $\mathbf{a}_{drag} = -\frac{1}{2}(C_D A_D/m)\rho \mathbf{v}_{rel} ^2 \hat{\mathbf{e}}_{v_{rel}}$ (2.6)
solar radiation pressure	STK truth	spherical spacecraft model, dual-cone eclipses
	propagator	spherical spacecraft model, no eclipses $\mathbf{a}_{SR} = -(p_{SR} C_{SR} A_{SR}/m) \hat{\mathbf{e}}_{r_{sat \rightarrow \odot}}$ (2.8)
secondary perturbations	STK truth	solid Earth tide, ocean tide, Earth albedo pressure, Earth thermal pressure
	propagator	none modelled

4. Perturbations tend to counteract each other. This is seen in **TOB-6**, which has a total error that is decidedly less than the sum of its parts.

2.4 Chapter Summary

In this chapter, we have introduced the concept of orbit propagation and demonstrated its reliance on accurate models of the accelerations acting on the propagated spacecraft. We have discussed a variety of primary and secondary perturbing accelerations, presenting mathematical models for the former and discounting the latter for the present study. We have compared the presented models to a known truth and found them sufficiently accurate; a summary of the propagator models and the truth can be found in Table 2.6. Having verified these models, we can now include them in our navigation studies ahead.

2.5 Propagator Configuration Tables

Table 2.7: Propagator Constants and Parameters

Ellipsoidal Constants (WGS-84)			
Earth gravitational parameter	μ_{\oplus}	398600.4418	km^3s^{-2}
radius of Earth	R_{\oplus}	6378.1370	km
Luni-Solar Constants (JPL DE-421)			
Sun gravitational parameter	μ_{\odot}	$132.7124400409 \times 10^9$	km^3s^{-2}
Moon gravitational parameter	μ_{ζ}	$4.902800076228 \times 10^3$	km^3s^{-2}
Drag Constants			
atmospheric density at 500 km [34]	ρ_{500}	4.89×10^{-13}	$\text{kg}\cdot\text{m}^{-3}$
drag surface area	A_D	10	m^2
satellite mass	m	1000	kg
drag coefficient	C_D	2.2	(unitless)
Earth angular velocity [26]	ω_{\oplus}	$7.29211585530 \times 10^{-5}$	$\text{rad}\cdot\text{s}^{-1}$
Solar Radiation Pressure (SRP) Constants			
solar radiation pressure at 1 AU (D)	p_{SR}	4.56×10^{-6}	$\text{N}\cdot\text{m}^{-2}$
SRP surface area	A_{SR}	10	m^2
satellite mass	m	1000	kg
SRP coefficient	C_{SR}	1.0	(unitless)

Table 2.8: STK 9.2.0 HPOP Force Models: Primary Perturbations

Earth Gravity	<ul style="list-style-type: none"> ▶ model: EGM-96 ▶ 70×70
Third-Body Gravity	<ul style="list-style-type: none"> ▶ model: point-mass ▶ Sun & Moon included
Drag	<ul style="list-style-type: none"> ▶ model: Jacchia-Roberts ▶ $F10.7_{daily} = F10.7_{avg} = 150.0$ ▶ $K_P = 3.0$ ▶ apparent Sun position
Solar Radiation Pressure	<ul style="list-style-type: none"> ▶ model: spherical ▶ apparent Sun position ▶ dual cone shadow ▶ lunar eclipse included ▶ Earth eclipse altitude: 0 km
Note: additional spacecraft parameters can be found in Table 2.7.	

Table 2.9: STK 9.2.0 HPOP Force Models: Secondary Perturbations

Solid Earth Tides	<ul style="list-style-type: none"> ▶ model: time dependent ▶ 70×70 ▶ minimum amplitude: 0 m
Ocean Tides	<ul style="list-style-type: none"> ▶ model: 30×30 ▶ minimum amplitude 0 m
Earth Albedo Pressure	<ul style="list-style-type: none"> ▶ $c_K = 1.0$ ▶ simple ground reflection
Earth Thermal Pressure	<ul style="list-style-type: none"> ▶ $c_K = 1.0$

Chapter 3

Measurements and their Models

3.1 Introduction

Measurements are essential to our understanding of the real world. Observations allow us to characterize a physical system and give us basis upon which to propose and test theories of system evolution. From the first blips of Sputnik to modern satellite ground stations, satellite operators have used measurements to determine their spacecraft state.

Measurement-based state determination of spacecraft is primary to the methods of propagation discussed in the previous chapter, for it is from measurements that these methods were developed. Centuries ago, Brahe's detailed observations of Mars allowed Kepler to discount the circular orbits of Copernicus and formulate his theory of elliptical motion [39]. More recently, ranging data from the Sputnik and Vanguard satellites were used to improve models of the Earth's gravitational field [28], a practice that continues to this day with satellites such as GRACE, GOCE, and CHAMP [40]. Measurements also form the basis of initial orbit determination; this process (which is beyond the scope of this thesis) provides the initial condition necessary for propagation as formalized in (2.1).

In the current study, we are looking to augment our predictions of satellite motion on HEO with measurements collected by the satellite as it orbits the Earth. These measurements will be functions of our spacecraft state and will thus contain information useful for our navigation solution. As discussed in Section 1.5.2, we will develop measurement models in the form of (1.4) – that is,

$$\mathbf{y} = \mathbf{h}(\mathbf{x}, t) \tag{3.1}$$

where measurements \mathbf{y} are obtained from state vector \mathbf{x} and time t through the vector function \mathbf{h} .

This chapter will first introduce a variety of measurements available to our HEO spacecraft. We will then examine certain measurements with potential to improve our autonomous navigation solution. This examination will detail a method of simulation for each measurement and present appropriate models $\mathbf{h}(\mathbf{x}, t)$ that represent the measurement and its errors. These models will allow us to use measurements to update our navigation solution, a process that will be discussed in Chapter 4.

Note on the simulation of measurements: As discussed in the introduction to this thesis (see Section 1.4), we are completing a navigation study through simulation. A necessary part of this simulation is the generation of measurements that mimic those that would actually be received by a spacecraft on our truth orbit. We will therefore make a distinction between two types of measurement model. Our *truth* measurement model is the model which we use to produce the measurements that are being provided to our navigation solution. Conversely, our *filter* measurement model is the model that the navigation solution uses internally to analyze the measurements it receives. Again, the latter process will be discussed more fully in Chapter 4.

3.2 Measurements for Autonomous Navigation

As discussed in Section 1.4, this study is concerned with the *autonomous* navigation of spacecraft in highly elliptical Earth orbits. We will therefore consider only those measurements that do not rely on the active participation of ground- or space-based assets. This precludes two common methods of spacecraft state determination: ranging (both radio and laser) and optical methods.

In his thesis, Hill [41] compiled a list of potential satellite-based measurements suitable for autonomous navigation. Included among these were a variety of truly *autonomous* sources: Sun- and star-trackers, horizon scanners (Earth or Moon), magnetometers, and pulsar-focused x-ray detectors. However, these measurements – which rely entirely on naturally observable phenomena – are limited in their ability to determine spacecraft state to the operational precision required for the current study.

Rather, spacecraft navigation has become reliant on so-called *semi-autonomous*

measurements – that is, measurements which rely on the pre-existence of some supporting ground- or space-based infrastructure. Under this category fall Global Navigation Satellite Systems (see Note below) and ground beacon systems (e.g., the French DORIS system). Though these measurements are not strictly autonomous, they do not require active participation from the supporting infrastructure – i.e., a spacecraft using these systems is only required to receive broadcast signals and does not itself need to communicate with any component of the system infrastructure. This preserves the autonomous nature of the navigation solution itself.

Falling neatly into neither of these categories is the inertial measurement unit (IMU), a collection of accelerometers and gyroscopes onboard a spacecraft that measures the magnitude and direction of non-conservative accelerations. Though not explicitly dependent on current spacecraft position, it can nevertheless provide information about the evolution of spacecraft position and velocity; see, for example, the work of Jah et al. [42] on aerobraking in the Martian atmosphere.

Building from the results of Hill [41] and others (see discussion in Section 1.3.1), we will examine GPS measurements in great detail. We will also discuss other measurements and examine their ability to augment our GPS-based solution.

Note on Global Navigation Satellite Systems: A global navigation satellite system (GNSS) is a constellation of satellites designed to provide navigation services to Earth-based users worldwide. At present, there are two such systems in full operation: GPS, maintained by the United States Air Force, and ГЛОНАСС / GLONASS¹, maintained by the Russian Aerospace Defence Forces. Two additional systems are currently under development: the European Galileo system and the Chinese BeiDou system²; both are expected to be operational by the end of the decade [43]. Despite this variety of GNSS options, this study focuses on the use of GPS alone. Of the four constellations, it is the most widely discussed in the literature, the most widely used in practise, and has a proven track record for space-based applications. Despite this focus, it is worth noting that the techniques and discussion developed in this study can be applied to measurements from any of the above systems.

¹Acronym for Глобальная навигационная спутниковая система / Globalnaya navigatsionnaya sputnikovaya sistema, literally Global Navigation Satellite System.

²Officially, the BeiDou Satellite Navigation System, pinyin *Bēidōu wèixīng dǎoháng xítōng*, from *bēidōu*, “northern dipper”.

3.3 The Global Positioning System: an introduction

The following section draws from an IEEE Spectrum article [44] written by Ivan Getting, the founding president of Aerospace Corp. and “the force behind the establishment of the Navstar Global Positioning System” [ibid.].

3.3.1 A brief history of American radionavigation systems

In the first half of the twentieth century, navigation technology leapt from star charts and chip logs to inertial navigation systems and radionavigation. The latter technique, radionavigation, employed the newly harnessed technology of radio waves to achieve accurate positioning in sunshine and cloud. An archetypical example was the ground-based LORAN (LOng RAnge Navigation) system; it featured coastal transmitters emitting synchronized signals. By measuring the differences in the time of arrival of these radio signals, shipborne users could determine their latitude and longitude (see Figure 3.1).

Following closely on the heels of LORAN was Transit, the first satellite-based navigation system. Inspired by the use of Doppler-shifted signals to reconstruct the orbit of Sputnik-1, Transit inverted this technique to allow users around the globe to determine their latitude and longitude based on the measured Doppler shift of the signal from a stable oscillator in a known orbit. Though Transit represented a major advancement from LORAN in both accuracy and coverage, it had its shortcomings; as discussed by Getting, it was involved, intermittent, and ill-suited for moving users.

In the early 1960s, the United States Air Force undertook initial studies to examine the navigation of aircraft using radio signals from satellites. System requirements necessitated continuous and all-weather 3-D passive positioning available to an unlimited number of users with relatively simple equipment. As the project developed, the name “Global Positioning System” was proposed – and thus, GPS was born.

3.3.2 GPS: Basis of Operation

At its most basic level, the Global Positioning System is a system of clocks. A spacecraft orbiting the Earth broadcasts a radio signal representing the accurate time

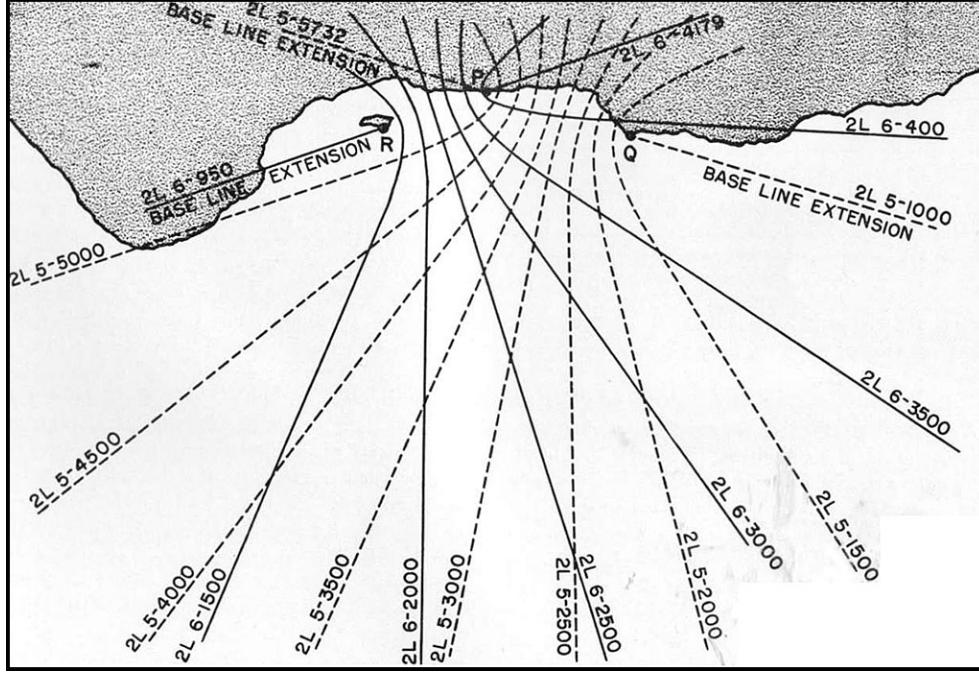


Figure 3.1: Families of hyperbolae from LORAN transmitters. P is a master station, Q and R are slave stations. Each hyperbola is a locus of all positions corresponding to one value for the time difference of arrival between the master and slave signals. A user with measurements for difference P-Q and difference P-R can be located at the intersection of the two corresponding hyperbolae. Public-domain figure courtesy of the U.S. Department of the Navy [45].

of an onboard clock (the *transmitter clock*). This time signal is received by a user on the ground, who compares it to a local clock (the *receiver clock*) and determines the offset between the two clocks. Multiplying by the constant speed of radio wave propagation converts this offset into the expected distance between the two clocks – a *pseudorange*. In other words, the time it took for the broadcast signal to reach the user is directly proportional to the distance between the satellite and the user. Assuming that the satellite’s position is well known, this pseudorange gives a measurement of the line-of-sight distance to a known (orbiting) control point – a measurement derived from the difference between two clocks.

3.3.3 GPS: System Architecture

The first GPS satellite was launched in 1978, with the system being declared fully operational in 1995 [46]. The operation of the Global Positioning System relies on

three key components: the space segment, the ground segment, and the user.

The **space segment** consists of a constellation of satellites in a twelve-hour, near-circular, medium Earth orbit (MEO) at an inclination of 55°. For global coverage, the System requires a minimum of twenty-four satellites distributed across six orbital planes. To ensure continuous operation, on-orbit spares are included in the constellation; as of December 2013, there were 31 active GPS satellites [47]. As noted above, each satellite has an onboard atomic clock. A signal representing this clock is broadcast to the Earth in the form of a repeating pseudo-random noise sequence; this sequence is unique to each broadcasting GPS satellite and allows the signals from different satellites to be differentiated. Modulated onto this sequence is a navigation message containing information necessary for the use of GPS signals in positioning, including parameters defining the satellite ephemeris and the correction required to match the onboard clock to GPS time³. For further details on the signal broadcast by the GPS constellation, please consult the GPS User Interface document [48].

The **ground segment** measures the orbits and clocks of the satellites, uploading the latest ephemerides and clock corrections to each satellite in the constellation. The clock corrections are referenced to GPS time. The ground segment thus ensures that the broadcast satellite orbit and clock data are sufficiently accurate for user positioning.

The **user** measures the clock signals received and compares them to the local receiver clock. A user with four satellites in view can measure four independent time offsets (yielding four pseudorange measurements) to solve for four unknowns: the user's three-dimensional position and user clock bias from GPS time. This calculation is reliant on the broadcast ephemerides and clock corrections. This carries the additional benefit of precise access to GPS time (and thus UTC).

3.4 GPS in this thesis: Models and Errors

As discussed in the introduction to this Chapter, we want to define measurement models as functions of state \mathbf{x} and time t . In this section, we will mathematically formulate the GPS pseudorange measurement and its various error sources. As introduced in Section 3.3.2, the pseudorange measurement approximates the straight-line

³GPS time is a monotonically increasing time scale identical to UTC with the notable omission of leap seconds.

distance between our spacecraft and the transmitting GPS satellite (hereinafter referred to as an SV, “space vehicle”). We will receive one pseudorange measurement for each SV for which our receiver has acquired lock⁴.

First, we will consider the *geometric pseudorange*, a measurement that relies only on geometry and assumes that the receiver clock is well-synchronized to GPS time. The measurement equation for geometric pseudorange is therefore:

$$\mathbf{h}_{GPS}(\mathbf{x}, t) = \text{col}_i [|\mathbf{r}(t) - \mathbf{r}_{SV(i)}(t_b)|] \quad (3.2)$$

where we define:

$\mathbf{h}_{GPS}(\mathbf{x}, t)$	geometric pseudorange measurement vector, km
$\mathbf{r}(t)$	position of our spacecraft at the time of signal reception, km
$\mathbf{r}_{SV(i)}(t_b)$	position of the i th GPS SV at the time of signal broadcast ⁵ , km

and col_i denotes a column vector where the i th row is given inside the brackets.

Equation (3.2) models an error-free measurement – the geometric pseudorange reported is a function of only spacecraft state (via position \mathbf{r}) and time (the position of each SV evolves with time). However, this equation – much like the two-body acceleration from the previous chapter – is but a starting point upon which we must build a more refined model. This model will include transmission errors, reception errors, and non-idealities in the propagation of the GPS signal itself. In the following sections, we will discuss these errors and present mathematical models where appropriate. This will enable us to deepen our understanding of the GPS pseudorange measurement, to better model the signals for simulation, and to better design an algorithm appropriate to the non-idealities of real-world GPS navigation.

Note on receiver clock error: When we introduced the GPS architecture (Section 3.3.3), we noted that four satellites are required for a GPS position lock. The fourth satellite is necessary to correct the difference between the receiver clock and GPS system time. The receiver clock bias can be a major error source if not handled

⁴For a further discussion of GPS lock, see Section 3.6.

⁵In a real GPS receiver, the SV position is evaluated at the estimated time of signal transmission. However, for computational simplicity this effect has been ignored in the current study; geometric pseudoranges are generated and estimated using the instantaneous position of the SV in question.

properly. However, as this clock bias can be treated as a state to be estimated rather than as a measurement error, it is not discussed in this section but is detailed in Section 3.5.

3.4.1 Tropospheric Error

The troposphere is the lowest region of the Earth’s atmosphere. Extending up to an altitude between 10 and 16 km, it contains four fifths of the mass of the atmosphere and virtually all of the atmosphere’s water content [49]. Radio-frequency signals (including those used by GPS) experience a delay when propagating through the troposphere due to variability of the refractive index of air under changing environmental conditions. Models for this delay have been developed for Earth-borne GPS users – see, for example, the work of Mendes and Langley [50]. These models can reduce resultant position errors to centimetres but require knowledge of local meteorological conditions.

In this study, we are concerned with the reception of GPS signals by a spacecraft in HEO. Our spacecraft has a perigee of approximately 500 km, well above the troposphere. Geometrically, very few of the GPS signals received will actually pass through the troposphere. This is demonstrated in the satellite visibility analysis in Figure 3.5. As such, in this study we will ignore any signal transiting through the troposphere by inflating the radius of the Earth by 100 km for the purpose of signal visibility calculations. This will also remove the need for local weather knowledge, simplifying our algorithm.

3.4.2 Ionospheric Error

The ionosphere is the region of the Earth’s atmosphere characterized by the presence of free ions, both positive and negative. The altitude of its extent varies greatly with solar activity, but it can generally be thought of as existing between 60 and 1000 km with a peak ion concentration at an altitude of 300 km [51]. These ions interact with radio signals transiting through the atmosphere, causing pseudorange signal delay. This has a direct effect on the accuracy of GPS positioning. As discussed by Klobuchar [52], this delay can reach 300 ns for ground-based users - translating into pseudorange errors of 100 m.

Fortunately, the ionospheric delay is frequency-dependent; a receiver capable of

receiving multiple GPS frequencies can remove it entirely. In this study, we make the worst-case assumption of a single-frequency GPS receiver processing only pseudorange measurements⁶. Thus, we will need to model the ionosphere twice: once for the generation of our measurements (in which we will perturb the ideal measurements) and once in our navigation algorithm.

Ionospheric delay - truth model: For the simulated measurements (our “truth” error), we adopted the model used by the Carleton GPS simulator (see Section 3.6.1 for simulator details). The model uses the work of Klobuchar [ibid.] as a basis. In brief, it calculates the vertical ionospheric delay predicted by Klobuchar’s model, which in turn relies on eight coefficients $\alpha_0\text{--}\alpha_3$ and $\beta_0\text{--}\beta_3$ transmitted with the GPS navigation message. The simulator then calculates an obliquity factor based on the current geometry between the satellite, the transmitting SV, and the Earth. This obliquity reflects the fact that the GPS signal is not transiting the ionosphere in a zenith-nadir sense.

The end result of this calculation is an equivalent pseudorange delay of up to 20 m at our perigee altitude of 500 km. This is added directly to the geometric pseudorange (3.2) calculated by the simulator. Mathematically, we can represent this as:

$$\mathbf{h}_{iono}^t(\mathbf{x}, t) = c \cdot \text{col}_i [\Delta T_{klob}(\mathbf{r}, \mathbf{r}_{SV(i)}, \alpha_j, \beta_j, t) \cdot OF(\mathbf{r}, \mathbf{r}_{SV(i)})] \quad (3.3)$$

where, in addition to the variables from (3.2), we have:

\mathbf{h}_{iono}^t	ionospheric delay truth model, km
c	speed of light, $299\,792.458\,\text{km}\cdot\text{s}^{-1}$
ΔT_{klob}	vertical ionospheric delay via Klobuchar model [52], km
α_j, β_j	ionosphere coefficients for Klobuchar model, unitless
OF	obliquity factor (current geometry) for Klobuchar model, unitless

and col_i once again denotes a column vector.

⁶Even with single-frequency receivers, it is possible to remove the ionosphere contribution. However, it requires the use of carrier phase measurements. For further details, please see Section 3.8.

Ionospheric delay - filter model: In our navigation algorithm, we again rely on the work of Klobuchar, whose algorithm has a demonstrated ability to remove 40–80% of the ionospheric delay. Using this as a baseline, we initialize our simulation with a randomly selected value from the range [0.4–0.8]. This is our *ionosphere factor*, IF , and represents the accuracy of the ionospheric model on a given day. Our filter model for ionospheric delay thus takes the form:

$$\mathbf{h}_{iono}^f(\mathbf{x}, t) = IF \cdot \mathbf{h}_{iono}^t(\mathbf{x}, t) \quad (3.4)$$

where \mathbf{h}_{iono}^f is the filter model for ionospheric delay, IF is the ionosphere factor described above, and \mathbf{h}_{iono}^t is calculated in (3.3). In short, our filter model removes a portion of the delay that our truth model introduces, mimicking the actual performance of the Klobuchar model for a single-frequency user. The final effect is therefore an adjustment to the measurements our filter receives:

$$\Delta \mathbf{y}_{iono} = (1 - IF) \cdot \mathbf{h}_{iono}^t(\mathbf{x}, t) \quad (3.5)$$

3.4.3 SV Ephemeris Errors

Broadcast along with the GPS navigation message are parameters defining the predicted orbits of the GPS satellites. These predictions are based on a four-hour curve fit and are updated every two hours (IS-GPS-200G [48], paragraph 20.3.4.4). Despite this frequent update, the broadcast orbits are subject to inaccuracies of a few metres. These errors in $\mathbf{r}_{SV(i)}$, the locations of our space-based control points, have a direct impact on the accuracy of our geometric pseudorange in (3.2).

In this thesis, we simulate these errors by purposefully using incorrect SV locations for the filter calculation of geometric pseudorange. The magnitude of these errors can be revealed by comparing the broadcast orbits to precise orbits obtained through post-processing⁷. The results of one such comparison can be found in Appendix E. With these results, we can model ephemeris errors by replacing $\mathbf{r}_{SV(i)}(t_b)$ in (3.2) with $\mathbf{r}_{SV(i)}^\Delta(t_b)$, which is defined as:

$$\mathbf{r}_{SV(i)}^\Delta(t_b) = \mathbf{r}_{SV(i)}(t_b) + [\Delta x_i \Delta y_i \Delta z_i]^T \quad (3.6)$$

where the components of the delta-vector can be found in Table E.1. Note that in

⁷Both broadcast and the precise orbits are available from the International GNSS Service [53].

this table, each satellite is referred to by its pseudo-random noise (PRN) number – analogous to the $SV(i)$ notation used in the equations above. We will therefore use a modified geometric pseudorange in our filter:

$$\mathbf{h}_{GPS}^{\Delta}(\hat{\mathbf{x}}, t) = \text{col}_i [|\hat{\mathbf{r}}(t) - \mathbf{r}_{SV(i)}^{\Delta}(t_b)|] \quad (3.7)$$

where $\hat{\mathbf{x}}$ reminds us that we are using the estimated spacecraft state (i.e., the filter's current estimate) and t_b denotes the time of broadcast for the GPS signal⁸.

While this constant offset model does not precisely reflect what occurs in real GPS operations (where the inaccuracy varies over the two hour interval between broadcast orbit updates), it tests the ability of our navigation algorithm to handle inaccuracies in the GPS ephemerides.

3.4.4 SV Clock Errors

As discussed in Section 3.3.2, GPS is fundamentally a system of clocks. Accurate GPS measurements are reliant on accurate clocks, where every nanosecond of clock error translates into approximately 30 cm of pseudorange error. The clocks onboard GPS satellites rely on highly stable rubidium and cesium oscillators, yet even these are subject to divergence from true GPS time [54].

Fortunately, these clock differences evolve in a predictable fashion. Though not entirely deterministic, they contain both periodic elements (where the fundamental period is approximately the twelve-hour orbit of the GPS satellites) and systematic clock variations [54]. The SV clock drifts are measured by the GPS ground segment, which calculates a second-order polynomial clock correction. This correction is then broadcast with the GPS navigation message. As with the SV ephemeris, the broadcast correction is updated every two hours.

From the GPS Interface Specification ([48], paragraph 20.3.3.3.3.1), the broadcast relationship between GPS time and the time of each SV clock is given by:

$$t_{SV(i)} = t_{GPS} + a_{f0} + a_{f1}(t_{GPS} - t_{oc}) + a_{f2}(t_{GPS} - t_{oc})^2 \quad (3.8)$$

where we define:

⁸See discussion in Footnote 5 (page 39).

$t_{SV(i)}$	effective clock time for i th SV at message transmission time, s
t_{GPS}	GPS system time, s
t_{oc}	clock data reference time, s
a_{fj}	clock correction coefficients for i th SV; s, $s \cdot s^{-1}$, $s \cdot s^{-2}$

In the previous section, we compared the broadcast SV ephemerides to precise ephemerides determined through post-processing. We can perform a similar comparison between the broadcast clock correction coefficients and their post-processed counterparts. The results of such an analysis can be found in Appendix E.

In this thesis, we will simulate errors in the SV clock-correction by perturbing the measurements received by our filter. Mathematically, we formulate this as:

$$\Delta \mathbf{y}_{SVclk} = c \cdot \text{col}_i [\Delta a_{f0,i} + \Delta a_{f1,i}(t_{GPS} \bmod 7200 \text{ s}) + \Delta a_{f2,i}(t_{GPS} \bmod 7200 \text{ s})^2] \quad (3.9)$$

where c is the speed of light, the values $\Delta a_{fj,i}$ can be found in Table E.2, and the 7200 s modulo operation represents the resetting of the clock-correction parameters every two hours. As in the previous section, each satellite is identified by its PRN number. Again, while this does not precisely model the true behaviour of clock-correction errors, it does test the ability of our navigation solution to handle realistic clock-correction inaccuracies.

3.5 GPS in this thesis: Receiver Clock Error

In the previous section, we discussed the reliance of GPS measurement accuracy on the accuracy of the SV clocks. The same argument applies to the clock inside the local GPS receiver. Unlike the highly accurate clocks onboard the SVs, GPS receivers typically use inexpensive clocks that are subject to significant divergence from GPS time. This divergence, known as *clock bias* and denoted by the variable b , is the difference between receiver clock time and GPS time:

$$b = t_{Rclk} - t_{GPS} \quad (3.10)$$

As discussed in Section 3.3.3, a user with four GPS satellites can solve for their

instantaneous position (x, y, z) and receiver clock bias b to within the limits of measurement error. GPS receivers exploit this to correct their own clock bias soon after acquiring the required signals. The clock can then be steered to (nearly) match GPS time, provided four satellites remain in view.

In the current study, we recognize that clock bias (a) has a direct effect on the pseudorange measurements we receive, and (b) is recoverable from said measurements. For this reason, we will treat b as a state and append it to our state vector as follows:

$$\mathbf{x}(t) = [r_x(t) \ r_y(t) \ r_z(t) \ v_x(t) \ v_y(t) \ v_z(t) \ b(t)]^T \quad (3.11)$$

Mathematically, we can model the measurement effect of receiver clock bias as:

$$\mathbf{h}_{Rclk}(\mathbf{x}, t) = c \cdot \text{col}_i[b] \quad (3.12)$$

where c is once again the speed of light. Since many GPS receivers remove their receiver bias automatically, measurement (3.12) is referred to as an *uncorrected* pseudorange. As highlighted by Moreau et al. [55], the processing (and modelling) of uncorrected pseudoranges ensures the observability of the receiver clock bias, an essential capability for navigation when less than four satellites are in view.

3.5.1 Propagating the Receiver Clock

In Chapter 2, we extensively discussed the means by which we propagate the position and velocity of our satellite. By modelling receiver clock error with its own state, we have created a corresponding need for receiver clock propagation.

To address the unique challenge of a HEO orbit, we adopted a hybrid model for receiver clock propagation. This model represents two distinct modes of clock operation expected for our receiver on HEO:

- a **steered** clock is constantly being adjusted to track GPS time within measurement error; while
- a **drifting** clock evolves according to stochastic processes without any adjustment by the receiver⁹.

⁹As discussed by Mikhailov and Vasil'ev [56], the deterministic clock models suited to the highly stable clocks onboard the GPS SVs are not suitable for less accurate receiver clocks. A stochastic model is thus the model of choice.

Clock-steering occurs when the GPS receiver has lock on at least four SVs (see Section 3.6) and the SVs have a favourable geometry relative to the receiver¹⁰. The completeness of the GPS constellation ensures that these conditions are almost always met for GPS receivers operating in LEO. However, the same cannot be said for our HEO satellite; GPS data will at times be insufficient for clock-bias correction.

We therefore require two different propagators: one for a steered receiver clock, the other for a drifting receiver clock. With these, we will generate a truth for our receiver clock bias (b_{true}), perturbing our simulated measurements accordingly:

$$\Delta \mathbf{y}_{Rclk} = c \cdot \text{col}_i[b_{true}] \quad (3.13)$$

As discussed more thoroughly in Chapter 4, our state estimator will also require a propagator. To retain the integrity of our navigation solution, we will purposefully downgrade its internal propagator. We will then include the estimated clock bias (b_{est}) in our navigation solution measurement model:

$$\mathbf{h}_{Rclk}(\hat{\mathbf{x}}, t) = c \cdot \text{col}_i[b_{est}] \quad (3.14)$$

Steered receiver clock: As discussed above, a GPS receiver with lock on at least four SVs will steer its receiver clock to match GPS time. This process of steering is internal to each receiver and its exact details are not necessarily known to the user. Thus, we must make an appropriate approximation based on observable data.

In a study completed for the German Aerospace Centre (DLR), Montenbruck [57] examined the behaviour of receiver clock error when the receiver clock is being steered. Based on his results ([ibid.], Figure 3.2) we adopted a coloured noise model for our steered clock:

$$\dot{b} = -\frac{1}{\tau_s} b + w_s \quad (3.15)$$

In the above, τ_s is our coloured noise decay constant and w_s is a zero-mean white noise process with standard deviation σ_s . Such a model is also known as a *first-order Gauss-Markov process* or an *exponentially correlated Markov process*.

Through trial and error we arrived at the parameters listed in Table 3.1, which provide a signal similar to that observed by Montenbruck. In the table, it is shown that the filter model for the steered clock is purposely corrupted. This reflects the

¹⁰Specifically, the Time Dilution of Precision (TDOP) must be below a set threshold of 10. For details on TDOP and its threshold, please see Appendix A.3.

real-world uncertainty in receiver performance.

Table 3.1: Steered receiver clock parameters

	Truth Model	Filter Model
τ_s	200 s	$200 \text{ s} + N(0, 20 \text{ s})$
σ_s	2 ns	$2 \text{ ns} + N(0, 0.2 \text{ ns})$

Drifting receiver clock: To accurately propagate our receiver clock when it is not being corrected, a more precise model of clock evolution is required. Dainty et al. [58] presented a stochastic model appropriate for ovenized crystal clocks typically used on spaceborne GPS receivers. The model is a two-state model, meaning that we propagate both clock bias b and its time derivative, clock drift $d = \dot{b}$. The model is defined as:

$$\dot{b} = d + w_b \quad (3.16a)$$

$$\dot{d} = w_d \quad (3.16b)$$

where noise vector $\mathbf{w}_k = [w_b \ w_d]^T$ is a zero-mean multivariate normal distribution with covariance \mathbf{Q}_{drift} . The latter is defined as:

$$\mathbf{Q}_{drift} = \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix} \quad (3.17)$$

where q_1 and q_2 are the bias and drift process noises for the ovenized crystal clock. The numerical values for both q_1 and q_2 can be found in Table 3.2.

Table 3.2: Drifting receiver clock parameters

	Truth Model	Filter Model
q_1	$1.6 \times 10^{-21} \text{ s}^2 \cdot \text{s}^{-1}$	$1.6 \times 10^{-21} \text{ s}^2 \cdot \text{s}^{-1} + N(0, 0.16 \times 10^{-21} \text{ s}^2 \cdot \text{s}^{-1})$
q_2	$3.2 \times 10^{-21} \text{ s}^2 \cdot \text{s}^{-3}$	$3.2 \times 10^{-21} \text{ s}^2 \cdot \text{s}^{-3} + N(0, 0.32 \times 10^{-21} \text{ s}^2 \cdot \text{s}^{-3})$

3.6 GPS in this thesis: Link Budget

In Section 3.4, we stated that our GPS receiver will record one pseudorange measurement for every SV for which the receiver has *lock*. This condition of lock is technical shorthand; in essence, it means that our GPS receiver is accurately receiving and decoding the navigation message transmitted by a given SV. Before lock can be achieved, two criteria must be met:

- first, there must be a clear line-of-sight vector between our receiver and the SV in question;
- second, the transmitted GPS signal must be strong enough to be successfully decoded by our receiver.

In the present study, we are simulating GPS pseudorange measurements with which to validate our navigation solution. A necessary part of this simulation is the mimicry of the SV lock condition; only when the above two criteria are met will our navigation solution receive the corresponding pseudorange measurement for processing.

The first criterion is a straightforward geometric calculation to ascertain whether the Earth¹¹ is blocking the line-of-sight vector between our receiver and the SV in question. The second criterion, however, is more subtle and requires the calculation of our situation-dependent *link budget*.

A link budget is a standard method of modelling the transmission of radio-frequency signals. It accounts for the gains and losses inherent to the transmitter, the receiver, and the medium of transmission. The summation of all gains and losses is then compared to a threshold figure which is typically specified by the manufacturer of the receiver¹². If the summation exceeds the threshold, the link budget has been met and the signal is considered valid.

In this study, we will use an adapted version of the link budget model presented by Wertz and Larson [34] (p. 551). For simplicity, we will ignore transmitter line losses and transmission path losses – the former are included in the transmitter power while the latter are zero in the vacuum of Earth orbit. The link budget thus takes the form:

¹¹and its atmosphere; see Section 3.4.1 for details.

¹²A more refined simulation would use separate thresholds for signal acquisition and signal tracking; see Weill [59] for further details and Potti et al. [60] for thresholds appropriate for HEO.

$$\frac{C}{N_0} = P_{tx} + G_{tx} + G_{rx} - L_{fs} - L_t \quad (3.18)$$

where we define:

C/N_0 ratio of received signal power to background noise¹³, dB·Hz

P_{tx} transmitter power, dB·W

G_{tx} transmitter antenna gain, dBi

G_{rx} receiver antenna gain, dBi

L_{fs} free space loss, dB

L_t temperature loss, dB·J

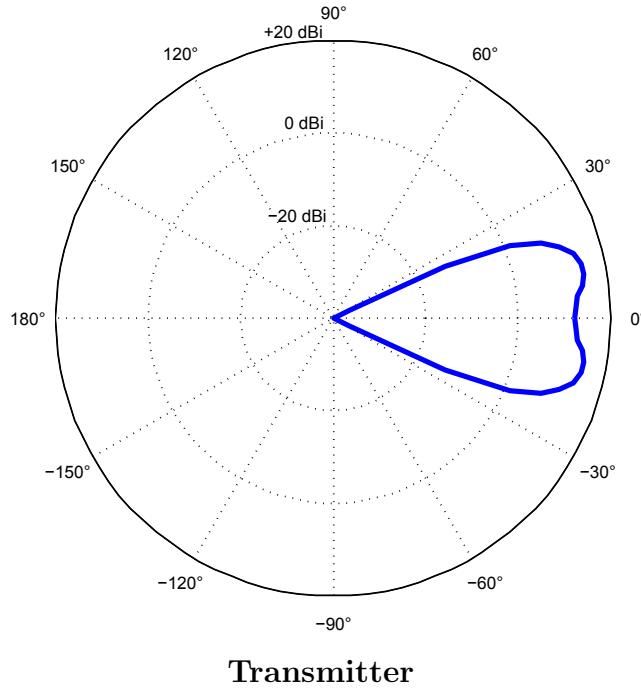
When generating our measurements, we will calculate the link budget and include the corresponding pseudorange if the result exceeds the C/N_0 threshold. That threshold, as well as the constant value for P_{tx} , can be found in Table 3.3. The remaining four parameters in (3.18) are explained below.

Table 3.3: Link budget constants. Parameters chosen to match those in a study completed by Magellan Aerospace [61] for direct comparison of results.

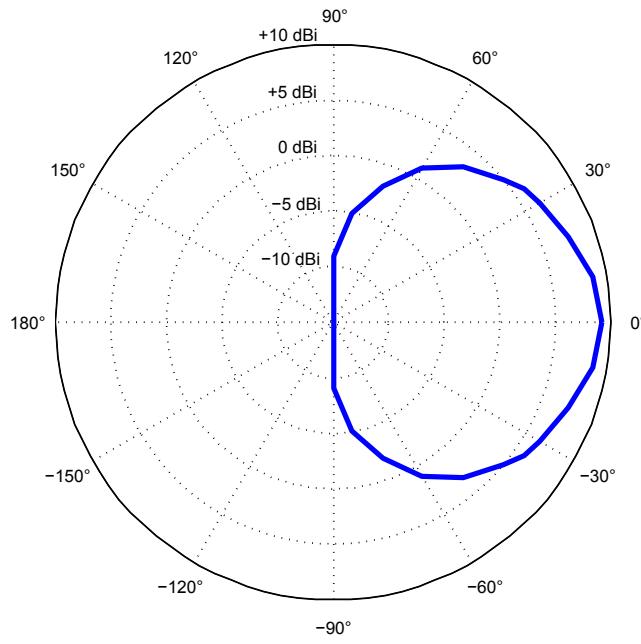
Parameter	Value	
C/N_0 threshold	35	dB·Hz
P_{tx}	10	dB·W
T_n	300	K

Antenna gains: The antenna gains G_{tx} and G_{rx} depend on both the attitude of our spacecraft and the relative geometry between our spacecraft and a given SV. The antenna patterns used in this study are axially symmetric with cross-sections given in Figure 3.2. The transmitter antenna aboard each SV is nadir-pointing, while the receiver antenna aboard our spacecraft can be configured to point towards zenith, nadir, or to evolve with the spacecraft attitude.

¹³More properly, the ratio of carrier power to white-noise spectral density.



Transmitter



Receiver

Figure 3.2: Cross-sectional gain of our axially symmetric antennae. Gain (dBi) is shown as a function of boresight angle. Patterns chosen to match those in a study completed by Magellan Aerospace [61] for direct comparison of results.

Temperature loss: The temperature loss L_t , expressed in dB·J, represents the thermal noise (the random vibration of electrons) inherent to any electrical system. It is a direct function of the equivalent noise temperature T_n of our receiver:

$$L_t = 10 \log k_B T_n \quad (3.19)$$

where the temperature-energy equivalence is realized by the Boltzmann constant ($k_B = 1.38 \times 10^{-23} \text{ J}\cdot\text{K}^{-1}$). The value for T_n used in this study can be found in Table 3.3.

Free space loss: The free space loss is an unavoidable fact of electromagnetic wave propagation. It results from the spherical spreading of electromagnetic radiation and is given by:

$$\begin{aligned} L_{fs} &= 10 \log \left[\left(\frac{4\pi df}{c} \right)^2 \right] \\ &= 20 \log \left(\frac{4\pi}{c} \right) + 20 \log f + 20 \log d \\ &= 20 \log d + 96.398 \text{ dB}\cdot\text{m}^{-1} \end{aligned} \quad (3.20)$$

where we use $d = |\mathbf{r}(t) - \mathbf{r}_{SV(i)}(t)|$ and have set $f = 1.575\,42 \text{ GHz}$, corresponding to the L1 frequency of the GPS system (the standard frequency for pseudorange measurements).

3.6.1 Calculating the link budget: the GPS simulator

A necessary requirement for the application of (3.2) is the knowledge of $\mathbf{r}_{SV(i)}(t)$ for all times of interest. To satisfy this requirement, this study employed the GPS simulator currently installed in the Department of Mechanical and Aerospace Engineering at Carleton University. The simulator is a LabPro-4000 manufactured by Navigation Laboratories Inc. of Laguna Niguel, CA, USA. When initialized with a truth trajectory (both translational and rotational) and properly configured, the simulator provides:

- $\mathbf{r}_{SV(i)}(t)$, the position of each SV as a function of time;
- geometric pseudorange (see Section 3.4);
- ionospheric delay (see Section 3.4.2); and

- attitude-dependent receiver antenna gain (see Section 3.6).

Further details about the operation of the GPS simulator can be found in Appendix C.1.

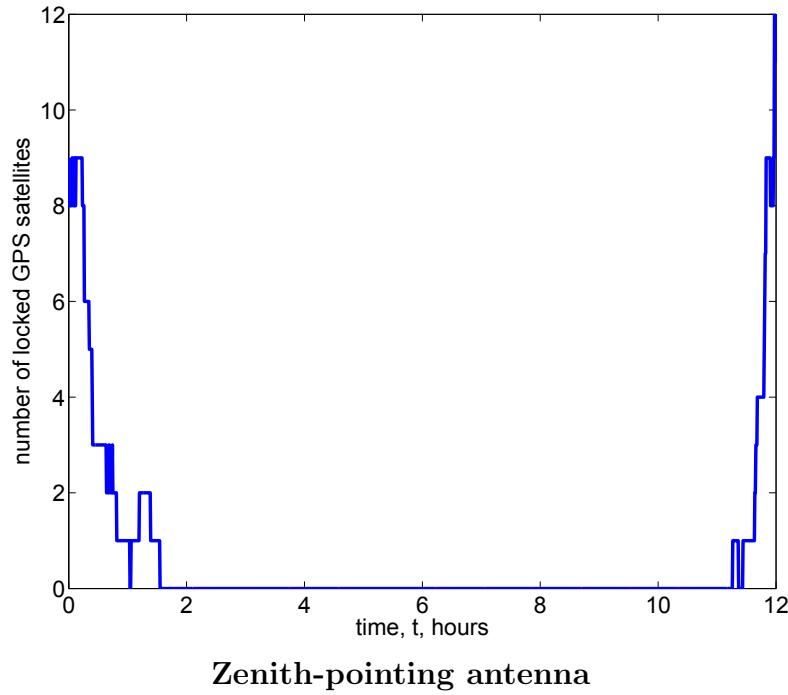
3.6.2 Applying the link budget: satellite visibility

Using the link budget outlined above, we combined our satellite truth trajectory with the SV ephemerides output from the GPS simulator (see Section 3.6.1) to examine our ability to lock onto SVs throughout our Molniya orbit.

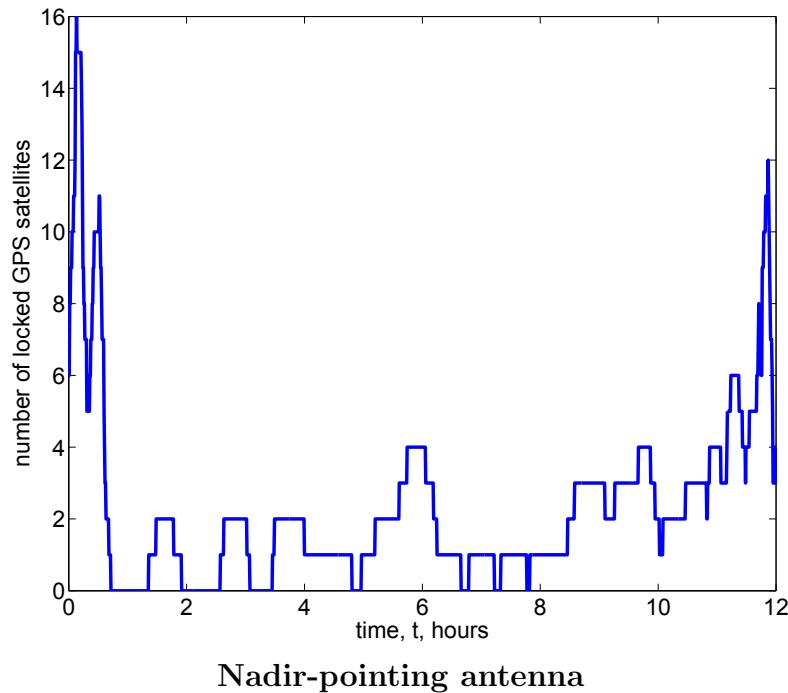
Nadir and Zenith visibility: To validate our link budget and SV ephemerides, we ran two cases: one where the receiver antenna was nadir pointing throughout the entire orbit, the other when it was zenith pointing. This was done for direct comparison with the Archimedes conceptual study completed by Potti et al. [60], which examined GPS visibility for a satellite in an 8 hour HEO. The visibility plots for our 12 hour orbit can be seen in Figure 3.3.

Consulting this figure, we can see that a zenith-pointing antenna has good signal reception at perigee, with eight or more locked SVs. However, this configuration experiences total outage over the entire apogee arc. This contrasts sharply with a nadir-pointing GPS antenna. This configuration has fewer locked SVs at perigee (as low as three), but quickly acquires more SVs as its altitude rises and maintains good SV visibility throughout the orbit. This ability to achieve lock at apogee is visualized in Figure 3.4, where a nadir-pointing antenna has a clear line of sight to an SV on the far side of the Earth. These conclusions match those from Figures 3–3 and 3–4 of Potti et al. [ibid.], strongly validating our link budget and SV ephemerides.

The effect of the troposphere: In our next analysis, we considered the effect of the troposphere on satellite visibility. As discussed in Section 3.4.1, we model the troposphere by ignoring any signal that passes within 100 km of the Earth’s surface. The impact of this decision on SV visibility can be seen in Figure 3.5, which shows the number of satellites blocked by our troposphere model over a single Molniya orbit. Though the tropospheric intersections are frequent – particularly near perigee – they are short-lived due to the rapidly changing orbital position of our spacecraft



Zenith-pointing antenna



Nadir-pointing antenna

Figure 3.3: GPS SV visibility over a single Molniya orbit for a single GPS antenna.
Perigee is located at 0 h and 12 h.

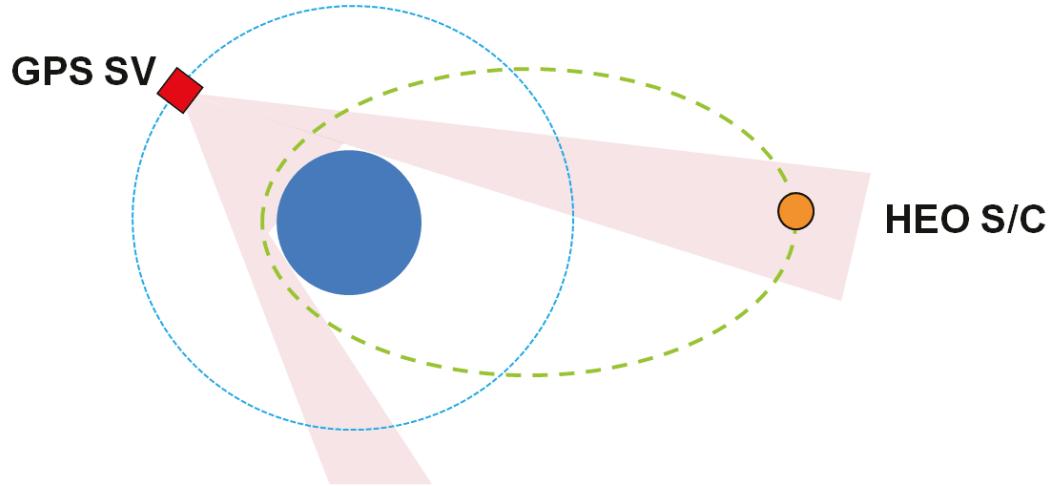


Figure 3.4: Visualization of GPS visibility on HEO. A zenith-pointing antenna will only receive GPS signals on the perigee arc when the spacecraft (HEO S/C) is below the altitude of the GPS constellation. A nadir-pointing antenna is able to lock on SVs on the opposite side of the Earth throughout the apogee arc. Figure is not to scale.

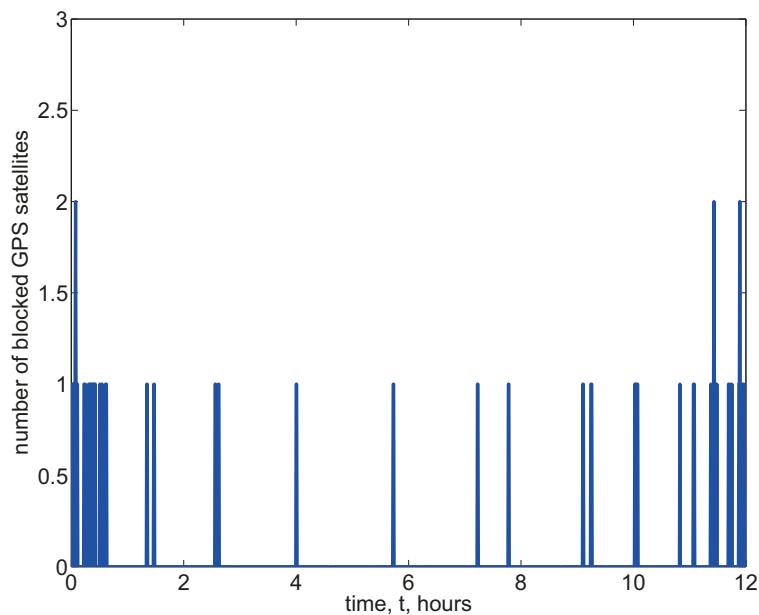


Figure 3.5: Impact of our tropospheric model on SV visibility over a single Molniya orbit. The blue spikes indicate times when a GPS signal was blocked by our 100 km thick atmosphere.

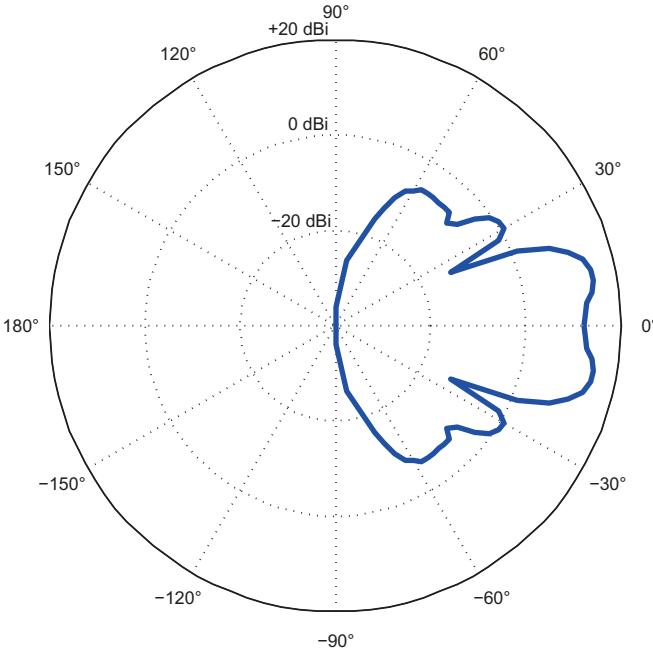


Figure 3.6: Cross-sectional gain of the GPS SV transmitter when side-lobes are included. Gain (dBi) is shown as a function of boresight angle.

and the SVs. Numerical analysis of this plot reveals that a signal passes through the troposphere only 5% of the time. Even when this occurs, it only occurs for one (rarely two) of the locked satellites at a time. We have therefore shown that our conservative model for the troposphere has a minimal impact on satellite visibility.

3.6.3 Refining the link budget: side-lobes

As visualized in Figure 3.2, the transmitting antenna onboard each SV is highly focused with a half-power beamwidth of 21.3° – 23.5° , depending on the year of SV manufacture [62]. This main beam is sufficient to cover the entire Earth from the altitude of the GPS constellation.

However, this is a simplified model. In accordance with antenna physics, the GPS transmitter antenna has side-lobes that surround the main lobe. A more realistic cross-section, which includes these side-lobes, can be seen in Figure 3.6.

In a study completed for the European Space Agency, Lorga [63] concluded that successful GPS-based navigation on HEO would require the reception of side-lobe signals. Such reception has been demonstrated in GEO [64] and in HEO [65]. It



Figure 3.7: The Canadian Space Agency ground station at SED Systems in Saskatoon, Saskatchewan. ©Daryl Mitchell [66]

is therefore natural for us include their effect in our study – an effect that will be explored more fully in Chapter 5.

3.7 Passive Ranging

In the previous sections, we have discussed GPS measurements as a well-suited means for autonomous navigation. In this section, we will introduce another measurement type: ground-based beacons. Ground-based beacons are a proven technology for satellite navigation – see, for example, the performance of the DORIS system noted in Section 3.8. However, navigation with such beacons has not yet been demonstrated beyond LEO. Thus, this study considers their inclusion in a manner consistent with the techniques and error sources developed for GPS, a proven technology for our highly elliptical orbit regime.

To maintain the spirit of autonomous navigation, our simulated beacons will be implemented to achieve *passive ranging*. They will continuously broadcast a signal that can be received by our satellite in the form of a pseudorange. For this scenario, we assume that two beacons are in operation at existing Canadian Space Agency ground stations in Saint-Hubert, Québec (45.48°N , 73.43°W) and Saskatoon, Saskatchewan (52.13°N , 106.68°W) (see Figure 3.7). It is further assumed that a high-quality dual-frequency GPS receiver would be co-located with the beacon. In addition to GPS measurements, this receiver would have access to corrections broadcast by the Global Differential GPS (GDGPS) system.

Having defined the scenario, we now turn our attention to the four main sources

of pseudorange error outlined in Section 3.4. For the purposes of passive ranging:

- as noted by Mendes and Langley [50], **tropospheric delay** can be reduced to an error level of centimetres using local weather measurements;
- **ionospheric delay** can be estimated using local ionospheric conditions – these conditions can be measured by the dual-frequency receiver as it monitors incoming GPS signals and their delay;
- **beacon location errors**, analogous to SV ephemeris errors, can be eliminated by a one-time determination of beacon location through a high-accuracy GPS survey;
- finally, **beacon clock errors**, analogous to SV clock errors, can be reduced to the order of nanoseconds using the dual-frequency receiver augmented with the GDGPS corrections (cf. Bar-Sever et al. [67]).

Having thus discounted the main sources of pseudorange error, we will apply a white-noise signal to model the remaining non-idealities in our system.

Beyond these considerations, these beacon signals are treated with the same amount of realism as the GPS signals. Valid signals must satisfy geometric considerations (specifically, an above-horizon constraint of 10°) and the link budget, including a transmitter profile identical to that of the GPS constellation. Simulated measurements will also be subject to receiver clock error derived from the same clock used in the GPS receiver. In effect, we have added two additional GPS satellites which are fixed to the surface of the Earth and broadcast outward. This was done in an attempt to minimize the additional hardware¹⁴ and software required for a passive ranging implementation.

Though such a system has not yet been implemented for HEO, the passive ranging scenario outlined above preserves the autonomous nature of our navigation solution – i.e., once the beacons have been constructed no active intervention on the part of the ground is required. Further, it is achievable with standard ground- and space-based components. We will explore the potential for passive ranging to improve the navigation solution in Section 5.4.5.

¹⁴ Additional receiver hardware would be required as our beacons would operate on a frequency different from that of the GPS constellation.

3.8 Other Measurements

In the preceding sections, we have presented GPS pseudorange measurements and passive ranging measurements as well-suited to our goal of autonomous spacecraft navigation. However, these are not the only measurements that might be available to our spacecraft. In this section, we will briefly discuss other measurements that were considered for use in our autonomous navigation solution.

Some of the measurements considered for this study have not yet been demonstrated beyond LEO:

- The NASA-led **Global Differential GPS (GDGPS)** network measures GPS signals from Earth and broadcasts corrections to land, sea, air, and space. These corrections all but eliminate ephemeris and ionospheric errors, enabling position accuracies of better than 30 cm (3D RMS) in LEO [68].
- The CNES¹⁵-led system of ground beacons, **DORIS** (Doppler Orbitography and Radiopositioning Integrated by Satellite), enables satellites to measure the Doppler shift of known transmitters. This system has been used by a number of LEO satellites for onboard real-time positioning, including the navigation of Jason-2 to better than 7.7 cm (RMS) in the radial component [69].
- The operational GLONASS¹⁶ constellation can be used to supplement GPS measurements. This application has been tested in space by various Roscosmos missions; notable among them is a Soyuz-Mir experiment that saw 3D 1σ position accuracy increase from 20 m (GPS alone) to 12 m (GLONASS/GPS) [70].

Should further studies or flight experiments demonstrate these systems beyond LEO, they each have potential to improve the autonomous navigation solution for a spacecraft in HEO.

Other measurements considered for this study are entirely feasible beyond LEO, yet they were not examined in greater detail due to time constraints:

- The most promising additional measurement is that of **GPS carrier phase**, a measurement derived from the GPS carrier signal. As noted by Montenbruck et al. [71], single frequency receivers can combine carrier phase measurements

¹⁵Centre National d'Études Spatiales, the space agency of France.

¹⁶See Note in Section 3.2

with pseudorange measurements to eliminate ionospheric error, enabling an accuracy of 1 m (3D RMS) in LEO.

- Another promising measurement is the use of an onboard **accelerometer** to update the spacecraft state during periods of high non-conservative acceleration (i.e. thrusting events or re-entry). As noted by Zanetti and D’Souza [72], such an accelerometer can minimize its own bias between thrusting events by constant self-calibration using other onboard measurements (e.g., GPS).
- A Magellan Aerospace study [73] under the STDP outlined in Section 1.3.3 examined the navigation of a spacecraft in a Tundra orbit (see Section 1.2.1) using purely **celestial measurements** – that is, images of the Earth, Moon, and stars using onboard star-trackers and imagers. Navigation solutions were achieved to within a few kilometres, agreeing with the findings of Hill [41].

After proper study has been complete, future satellites could fuse one or more of these additional measurements as a means to overcome some of the GPS-based positioning inaccuracies noted throughout this chapter.

3.9 Chapter Summary

In this chapter, we have introduced the concept of measurements and outlined the role they will serve in our autonomous navigation solution. First, we detailed the GPS pseudorange measurement, introducing models for GPS signal propagation and models for the major pseudorange error sources (see Table 3.4). Related to this measurement, we introduced two new states modelling our receiver clock and detailed a means of simulating its impact on our psuedorange measurements. Next, we introduced the passive ranging measurement, a beacon-based form of ranging that is compatible with our goal of autonomous navigation. Finally, we outlined a number of measurements not considered in this study and summarized their potential to improve our navigation solution. With these tools in hand, we are prepared to incorporate measurements in our autonomous navigation solution.

Table 3.4: Summary of measurement errors and models for GPS pseudorange

tropospheric delay	truth	signals passing within 100 km are blocked
	filter	not modelled
ionospheric delay	truth	measurements perturbed by scaled Klobuchar model $\Delta \mathbf{y}_{iono} = (1 - IF) \cdot \mathbf{h}_{iono}^t(\mathbf{x}, t)$ (3.5)
	filter	not modelled
SV ephemeris error	truth	uses exact ephemerides $\mathbf{h}_{GPS}(\mathbf{x}, t) = \text{col}_i [\mathbf{r}(t) - \mathbf{r}_{SV(i)}(t)]$ (3.2)
	filter	uses perturbed ephemerides $\mathbf{h}_{GPS}^\Delta(\hat{\mathbf{x}}, t) = \text{col}_i [\hat{\mathbf{r}}(t) - \mathbf{r}_{SV(i)}^\Delta(t)]$ (3.7)
SV clock error	truth	measurements perturbed by time-varying clock error $\Delta \mathbf{y}_{SVclk} = c \cdot \text{col}_i [\Delta a_{f0,i} + \Delta a_{f1,i}(\diamond) + \Delta a_{f2,i}(\diamond)^2]$ (3.9) $(\diamond = t_{GPS} \bmod 7200 \text{ s})$
	filter	not modelled
receiver clock error	truth	measurements perturbed by clock bias truth $\Delta \mathbf{y}_{Rclk} = c \cdot \text{col}_i [b_{true}]$ (3.13)
	filter	filter estimates clock bias and its effect $\mathbf{h}_{Rclk}(\hat{\mathbf{x}}, t) = c \cdot \text{col}_i [b_{est}]$ (3.14)
model summary	truth	$\mathbf{y} = \mathbf{h}_{GPS}(\mathbf{x}, t) + \Delta \mathbf{y}_{iono} + \Delta \mathbf{y}_{SVclk} + \Delta \mathbf{y}_{Rclk}$
	filter	$\hat{\mathbf{y}} = \mathbf{h}_{GPS}^\Delta(\hat{\mathbf{x}}, t) + \mathbf{h}_{Rclk}(\hat{\mathbf{x}}, t)$

Chapter 4

Nonlinear Filter Theory

In the context of navigation, filtering is the fusion of propagation and measurement to achieve a more accurate solution than either method can provide alone. Returning to the age of sail, the centuries-old art of dead reckoning was greatly enhanced by the updated position estimates permitted by the marine chronometer and sextant. We will echo this example of fusion for navigation as we combine our dynamic and measurement models from Chapters 2 and 3 in our quest for the optimal navigation solution for a spacecraft on HEO.

As introduced in Section 1.5.3, the classic algorithm used for filtering of dynamic systems is the Kalman Filter. In the words of the National Academy of Engineering¹, “The Kalman filter revolutionized the field of control theory and has become pervasive in engineering systems.” [74] Among these many systems are those designed for inertial navigation, for airborne submarine detection, and for radar tracking of the now-retired Space Shuttle [ibid.].

The Kalman filter owes its ubiquity to a number of key factors. First, it is a recursive filter; once the dynamic and measurement models are defined, the current state estimate (and corresponding estimate of state uncertainty) is calculated based only on the preceding estimate and a current measurement. This recursive nature and the corresponding computational simplicity contrasts sharply with batch filtering methods (i.e. a least squares curve-fitting). Second, despite its recursive nature, each estimate is reliant on all previous estimates and measurements – these are captured in the state uncertainty which is carried forward through each estimation. Finally, the Kalman filter contains explicit variables defining system uncertainty; these can be adjusted (“tuned”) for optimal filter performance.

¹Citation for the 2008 Draper Prize.

In this chapter, we will first introduce Kalman's linear filter. Next, we will outline a variety of nonlinear filters appropriate to our non-linear problem of satellite navigation. Finally, we will give a brief sampling of the many missions to date which have used a Kalman filter for navigation. With this basis established, we will be ready for the application of these filters in Chapter 5.

4.1 The Discrete Linear Kalman Filter

4.1.1 Problem Definition

Following the derivation of de Ruiter et al. [10] (p. 477), we shall examine a system of the form²:

$$\mathbf{x}_k = \Phi_{k-1} \mathbf{x}_{k-1} + \mathbf{w}_{k-1} \quad (4.1a)$$

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \quad (4.1b)$$

where we define:

\mathbf{x}_k state vector, $\mathbf{x}_k \in \mathbb{R}^n$

Φ_k state transition matrix, $\Phi_k \in \mathbb{R}^{n \times n}$

\mathbf{w}_k modelling error, $\mathbf{w}_k \in \mathbb{R}^n$

\mathbf{y}_k measurement vector, $\mathbf{y}_k \in \mathbb{R}^m$

\mathbf{H}_k measurement model matrix, $\mathbf{H}_k \in \mathbb{R}^{m \times n}$

\mathbf{v}_k measurement error, $\mathbf{v}_k \in \mathbb{R}^m$

and the subscript k refers to time t_k . We will treat \mathbf{x}_k , \mathbf{w}_k , \mathbf{y}_k , and \mathbf{v}_k as random variables. We will also assume that \mathbf{x}_k , \mathbf{w}_k , and \mathbf{v}_k are jointly independent for all t_k . Finally, we will assume that our error vectors obey:

$$\begin{aligned} E(\mathbf{w}_k) &= \mathbf{0} & E(\mathbf{w}_k \mathbf{w}_j^T) &= \mathbf{Q}_k \delta_{kj} > \mathbf{0} \\ E(\mathbf{v}_k) &= \mathbf{0} & E(\mathbf{v}_k \mathbf{v}_j^T) &= \mathbf{R}_k \delta_{kj} \geq \mathbf{0} \end{aligned} \quad (4.2)$$

²Note that this system has no input term. This is due to two factors. First, the system considered in this study is not controlled. Second, the inclusion of a control term would not affect the performance of the filter as the control term does not appear in the estimation error dynamics.

where $\mathbf{Q}_k \in \mathbb{R}^{n \times n}$ is the process noise covariance, $\mathbf{R}_k \in \mathbb{R}^{m \times m}$ is the measurement noise covariance, and δ_{kj} is the Kronecker delta.

The Kalman filter is a recursive method to calculate state estimate $\hat{\mathbf{x}}_k$ at time t_k . For each recursion, it requires a previous state estimate $\hat{\mathbf{x}}_{k-1}$, the corresponding state covariance:

$$\mathbf{P}_{k-1} = E \left((\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1})(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1})^T \right), \quad (4.3)$$

and a measurement \mathbf{y}_k . From these, the Kalman filter will provide a state estimate $\hat{\mathbf{x}}_k$ that “best” approximates the true state \mathbf{x}_k to within the limits of the noise vectors (i.e., to within the limits defined by \mathbf{Q}_k and \mathbf{R}_k). This “best” approximation is defined as the estimate which minimizes the sum of the squares of the error in each component of our state vector:

$$J = (\mathbf{x}_k - \hat{\mathbf{x}}_k)^T (\mathbf{x}_k - \hat{\mathbf{x}}_k) = \text{trace}(\mathbf{P}_k) \quad (4.4)$$

where J is the cost function which we are looking to minimize. For this reason, the Kalman filter is also known as a *minimum variance estimator*.

4.1.2 Algorithm

The Kalman filter algorithm can be understood as a two-part process:

- the **time update**, where the physical model of the system as defined in (4.1a) is used to propagate the solution from time t_{k-1} to time t_k . The resulting state estimate $\hat{\mathbf{x}}_k^-$ and covariance \mathbf{P}_k^- are denoted *a priori* (Latin, “from the former”).
- the **measurement update**, where measurements of the system state as defined in (4.1b) are used to update the solution at time t_k . The resulting state estimate $\hat{\mathbf{x}}_k^+$ and covariance \mathbf{P}_k^+ are denoted *a posteriori* (Latin, “from the latter”).

With these definitions in mind, the algorithm proper takes the form:

The Kalman Filter

Time update:

$$\hat{\mathbf{x}}_k^- = \Phi_{k-1} \hat{\mathbf{x}}_{k-1}^+ \quad (4.5a)$$

$$\mathbf{P}_k^- = \Phi_{k-1} \mathbf{P}_{k-1}^+ \Phi_{k-1}^T + \mathbf{Q}_{k-1} \quad (4.5b)$$

Measurement update:

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (4.5c)$$

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-) \quad (4.5d)$$

$$\mathbf{P}_k^+ = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^- (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \quad (4.5e)$$

where \mathbf{K}_k is the Kalman gain (discussed below), \mathbf{I} is an appropriately sized identity matrix, and all other variables retain their definitions from the previous section.

The genius of the Kalman filter lies in the measurement correction in (4.5d). In this equation, we compute the difference between our true measurement \mathbf{y}_k and our predicted measurement $\hat{\mathbf{y}}_k = \mathbf{H}_k \hat{\mathbf{x}}_k^-$. This difference is known as the *innovation* or *measurement residual*. The state estimate $\hat{\mathbf{x}}_k^-$ is then corrected by this innovation through some gain \mathbf{K}_k . This so-called Kalman gain, calculated in (4.5c), is designed specifically to minimize the cost function in (4.4).

The Kalman filter is thus at its most fundamental form a predictor-corrector; it predicts the a priori state using the physical model, then updates this estimate with measurements according to the measurement model. The minimization of state variance, combined with the form of the estimator assumed by (4.5a) and (4.5d), results in the Kalman filter's status as the optimal *linear* estimator – that is, the optimal estimator using only linear functions.

For further information, please consult Kalman's original paper [22] or an appropriate textbook (e.g. Crassidis and Junkins [75], Simon [76]).

4.2 Nonlinear Kalman Filters

As discussed in the previous section, the Kalman Filter is the optimal linear estimator. However, the majority of real-world systems are nonlinear – including all models introduced in Chapters 2 and 3. We recall that such systems can be modelled as:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t) + \mathbf{w}(t) \quad (4.6a)$$

$$\mathbf{y}(t_k) = \mathbf{h}(\mathbf{x}(t_k), t_k) + \mathbf{v}(t_k) \quad (4.6b)$$

where all variables retain their definitions from the previous section. Note that while the dynamic model is continuous on all times t , the measurement model is evaluated only at discrete times t_k .

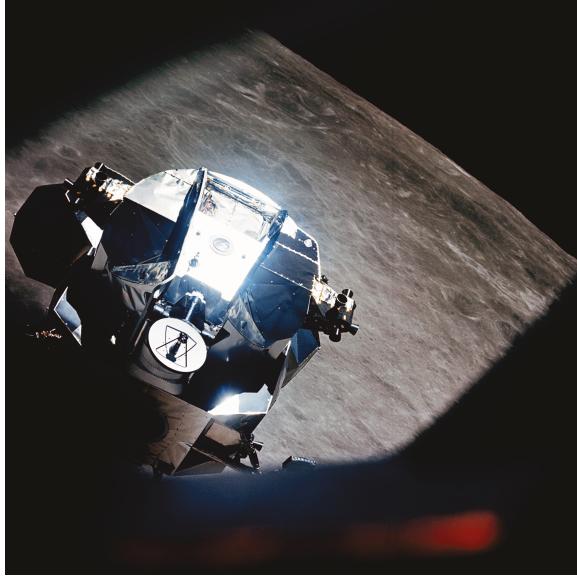


Figure 4.1: Apollo 10 Lunar Module *Snoopy* as seen from Command Module *Charlie Brown* shortly before rendezvous in lunar orbit. A Kalman filter onboard *Snoopy* converted radar and IMU measurements into the precise relative navigation solution required for a successful rendezvous [13].
Public-domain image courtesy of NASA [77].

We are therefore seeking an algorithm similar to the Kalman Filter which is appropriate to system (4.6). In the following sections, three such algorithms will be introduced, each with their own strategy to apply the optimal linear estimator to the general nonlinear system.

4.2.1 Extended Kalman Filter

The following section draws from an IEEE Control Systems Magazine article [74] written by Mohinder Grewal and Angus Andrews, coauthors of a number of works on the Kalman filter and its application.

Shortly after Kalman published his seminal paper, he was invited to present his results at the NASA Ames Research Centre (ARC) in Mountain View, California. The Dynamics Analysis Branch at ARC had been tasked with a monumental challenge: trajectory estimation and control for a manned mission to the Moon. Kalman's recursive solution intrigued the Branch chief, who initiated a campaign of intense

simulation on the ARC mainframe to demonstrate the validity of the Kalman filter through repeated Monte Carlo simulations.

Having verified the Kalman filter, the engineers at ARC sought a means by which it could be applied to the non-linear problem of spacecraft trajectory estimation. Their solution to this problem came to be known as the *Extended Kalman Filter* (EKF). Further Monte Carlo simulations demonstrated the ability of the EKF to retain sufficient accuracy for the Apollo trajectory problem on a mere 36 bits of floating point arithmetic. The rest, as highlighted in Figure 4.1, is history.

In the decades since this successful implementation, the EKF has become the *de facto* standard for nonlinear state estimation. Its fundamental simplification is a first-order Taylor expansion of the functions $\mathbf{f}(\mathbf{x}, t)$ and $\mathbf{h}(\mathbf{x}, t)$ about a reference trajectory which follows the noise-free dynamics ($\mathbf{w} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$, as appropriate). This Taylor expansion is truncated after the first order. We therefore require the Jacobians of our nonlinear models:

$$\mathbf{F}(\mathbf{x}(t), t) = \frac{\partial}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}, t) \quad (4.7a)$$

$$\mathbf{H}(\mathbf{x}(t), t) = \frac{\partial}{\partial \mathbf{x}} \mathbf{h}(\mathbf{x}, t) \quad (4.7b)$$

Note that while these functions $\mathbf{F}(\mathbf{x}, t)$ and $\mathbf{H}(\mathbf{x}, t)$ are notationally similar to the matrices in (4.1), they are distinct in their definition and role.

The algorithm itself is outlined below. Note that in the interest of concise notation, we have adopted the convention $\mathbf{x}(t_k)^- \Rightarrow \mathbf{x}_k^-$. As with the Kalman filter, we seek an estimate $\hat{\mathbf{x}}_k^+$ and its covariance \mathbf{P}_k^+ . These quantities are obtained from their previous values $(\hat{\mathbf{x}}_{k-1}^+, \mathbf{P}_{k-1}^+)$ and are improved by measurements \mathbf{y}_k .

The Extended Kalman Filter

Time update:

$$\hat{\mathbf{x}}_k^- = \hat{\mathbf{x}}_{k-1}^+ + \int_{t_{k-1}}^{t_k} \mathbf{f}(\hat{\mathbf{x}}(\tau), \tau) d\tau \Big|_{\hat{\mathbf{x}}(t_{k-1})=\hat{\mathbf{x}}_{k-1}^+} \quad (4.8a)$$

$$\Phi(t_k, t_{k-1}) = \exp [\mathbf{F}(\hat{\mathbf{x}}_{k-1}^+, t_{k-1}) \cdot (t_k - t_{k-1})] \quad (4.8b)$$

$$\mathbf{P}_k^- = \Phi(t_k, t_{k-1}) \mathbf{P}_{k-1}^+ \Phi(t_k, t_{k-1})^T + \mathbf{Q}_{k-1} \quad (4.8c)$$

Measurement update:

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}(\hat{\mathbf{x}}_k^-, t_k)^T [\mathbf{H}(\hat{\mathbf{x}}_k^-, t_k) \mathbf{P}_k^- \mathbf{H}(\hat{\mathbf{x}}_k^-, t_k)^T + \mathbf{R}_k]^{-1} \quad (4.8d)$$

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k [\mathbf{y}_k - \mathbf{h}(\hat{\mathbf{x}}_k^-, t_k)] \quad (4.8e)$$

$$\mathbf{P}_k^+ = [\mathbf{I} - \mathbf{K}_k \mathbf{H}(\hat{\mathbf{x}}_k^-, t_k)] \mathbf{P}_k^- [\mathbf{I} - \mathbf{K}_k \mathbf{H}(\hat{\mathbf{x}}_k^-, t_k)]^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \quad (4.8f)$$

In the above, we note that our Jacobians $\mathbf{F}(\mathbf{x}, t)$ and $\mathbf{H}(\mathbf{x}, t)$ appear explicitly in the calculations of the covariance matrices and Kalman gain; this is indicative of the linearization that occurs during the derivation of the EKF. However, the full nonlinear models are retained for the state propagation (4.8a) and measurement update (4.8e); this improves the accuracy of our state estimates without abandoning the form of the Kalman filter.

Further examining (4.8a), it appears that the EKF algorithm calls for the integration of the random variable $\hat{\mathbf{x}}(t)$ through the function $\mathbf{f}(\hat{\mathbf{x}}(t), t)$. Though the physical model in (4.6a) is stochastic and the estimate $\hat{\mathbf{x}}(t)$ is itself based on the random measurements given in (4.6b), the EKF works with realizations of both the measurements and the state estimate. The numerical integration in the EKF is therefore deterministic and is based on the Riemann definition of the integral.

In the case of the time update for this particular formulation of the EKF, a further assumption has been made. As a result of our linearization, the state transition matrix $\Phi(t_k, t_{k-1})$ now corresponds to the differential equation:

$$\frac{d}{dt} \tilde{\mathbf{x}} = \mathbf{F}(\hat{\mathbf{x}}(t), t) \tilde{\mathbf{x}} \quad (4.9)$$

where $\tilde{\mathbf{x}} = \mathbf{x} - \hat{\mathbf{x}}$ is the error in our state estimate. This equation is posed on the interval $[t_{k-1}, t_k]$ with initial condition $\tilde{\mathbf{x}}_{k-1}^+ = \mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}^+$. For sufficiently small sample periods we may treat $\mathbf{F}(\hat{\mathbf{x}}(t), t)$ as a constant. The state transition matrix corresponding to such an equation is the one given in (4.8b).

A final clarification can be made regarding the process noise covariance \mathbf{Q}_{k-1} . Recalling the discrete-time definition of \mathbf{Q}_{k-1} in (4.2), we can define its continuous-time counterpart as:

$$E(\mathbf{w}(t)\mathbf{w}(\tau)^T) = \mathbf{Q}(t)\delta(t - \tau) \quad (4.10)$$

where $\delta(t - \tau)$ is the Dirac delta function. \mathbf{Q}_{k-1} is then obtained by integrating $\mathbf{Q}(t)$

over the period of the time update:

$$\mathbf{Q}_{k-1} = \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau) \mathbf{Q}(\tau) \Phi(t_k, \tau)^T d\tau \quad (4.11)$$

allowing us to incorporate knowledge of our continuous-time process noise $\mathbf{w}(t)$ into the EKF algorithm.

In summary, the Extended Kalman Filter is a straightforward means by which to apply the Kalman filter algorithm to a non-linear system. It has successfully been applied to many different estimation problems, including the Apollo trajectory estimation problem. However, its fundamental basis on a linearization about a reference trajectory requires that estimation errors remain small for accurate filter operation.

4.2.2 Unscented Kalman Filter

As outlined in the previous section, the Extended Kalman Filter allows us to apply the Kalman Filter algorithm to a nonlinear system by means of a linearization about a reference trajectory. However, this linearization has its limitations. Determination of the optimal values for process noise \mathbf{Q}_{k-1} and measurement noise \mathbf{R}_k (i.e., *tuning the filter*) can be a time-consuming task. Even a well-tuned filter can experience difficulties if operational conditions cause the nonlinearities of a system to dominate the system dynamics. Such was the case for the Apollo docking EKF, which made heavy use of an “astronaut-in-the-loop” and required frequent restarting [78].

Learning from these experiences, analysts in the latter part of the 20th century began to reformulate the Kalman filter for application to nonlinear systems. One such reformulation is the Unscented Kalman Filter (UKF), first proposed by Julier and Uhlmann [79]. The UKF retains the Kalman filter’s elegant architecture: state, process, and measurement uncertainties are all captured through covariance matrices. However, the UKF avoids the need for linearization (i.e., the need for Jacobians) by means of the unscented transform. For this and other reasons, the UKF has gained widespread acceptance; notably, it played a key role during the entry, descent, and landing sequence for the Mars Science Laboratory [80].

In the rest of this section, we will define the unscented transform and present an algorithm for the UKF, a filter which incorporates this transform twice: once in the time update and again in the measurement update. It is worth noting that we will be using the *scaled unscented transform*; as outlined by Wan and van der Merwe [81], the

additional parameters of this transform allow us to optimize the unscented transform for a given application. These parameters will be further discussed in Section 5.5.

Note on the word “unscented”: In an interview, Uhlmann [78] explained the naming of his proposed transform and filter:

One evening everyone else in the lab was at the Royal Opera House, and as I was working I noticed someone’s deodorant on a desk. The word “unscented” caught my eye as the perfect technical term. At first people in the lab thought it was absurd... [yet] within a few months we had a speaker visit from another university who talked about his work with the “unscented filter.”

The Unscented Transform

To introduce the unscented transform, we shall consider the transformation of state estimate $\hat{\mathbf{x}} \in \mathbb{R}^n$ and state covariance $\mathbf{P} \in \mathbb{R}^{n \times n}$ through a nonlinear function $\mathbf{z} = \mathbf{g}(\mathbf{x}) \in \mathbb{R}^m$. The unscented transform replaces $\hat{\mathbf{x}}$ and \mathbf{P} with a set of sigma points \mathcal{X}_i where $i = 0, \dots, 2n$. These sigma points are calculated directly from $\hat{\mathbf{x}}$ and \mathbf{P} and are defined such that:

$$\hat{\mathbf{x}} = \sum_{i=0}^{2n} w_i^m \mathcal{X}_i \quad (4.12)$$

$$\mathbf{P} = \sum_{i=0}^{2n} w_i^c (\mathcal{X}_i - \hat{\mathbf{x}})(\mathcal{X}_i - \hat{\mathbf{x}})^T \quad (4.13)$$

where w_i^m and w_i^c are appropriately defined mean and covariance weights. Next, we pass each of our sigma points through our nonlinear function $\mathbf{g}(\mathbf{x})$ to obtain our transformed sigma points \mathcal{G}_i :

$$\mathcal{G}_i = \mathbf{g}(\mathcal{X}_i) \quad i = 0, \dots, 2n \quad (4.14)$$

Finally, we calculate the mean and covariance of our transformed sigma points:

$$\hat{\mathbf{z}} = \sum_{i=0}^{2m} w_i^m \mathcal{G}_i \quad (4.15)$$

$$\mathbf{P}_z = \sum_{i=0}^{2m} w_i^c (\mathcal{G}_i - \hat{\mathbf{z}})(\mathcal{G}_i - \hat{\mathbf{z}})^T \quad (4.16)$$

Thus, we can obtain an approximation to our transformed state $\hat{\mathbf{z}}$ and its covariance \mathbf{P}_z without the need to evaluate the Jacobian of our nonlinear function $\mathbf{g}(\mathbf{x})$.

The Scaled Unscented Kalman Filter

With the unscented transform in hand, we can define our UKF algorithm as follows. Given $\hat{\mathbf{x}}_{k-1}^+$, \mathbf{P}_{k-1}^+ , and a measurement $\mathbf{y}_k \in \mathbb{R}^m$:

1. Choose scaled unscented transform parameters α , κ , and β , and calculate:

$$\lambda = \alpha^2(n + \kappa) - n \quad (4.17a)$$

where λ is a scaling parameter and n is the dimension of \mathbf{x} .

2. Perform the upper Cholesky decomposition (analogous to a matrix square root) of the state covariance, $\mathbf{S}_{k-1}^+ = \text{chol}(\mathbf{P}_{k-1}^+)$, and calculate the sigma points:

$$\begin{aligned} \mathcal{X}_{0,k-1}^+ &= \hat{\mathbf{x}}_{k-1}^+, \\ \mathcal{X}_{i,k-1}^+ &= \hat{\mathbf{x}}_{k-1}^+ + \sqrt{n + \lambda} \cdot \mathbf{s}_{k-1,i}^+, \\ \mathcal{X}_{i+n,k-1}^+ &= \hat{\mathbf{x}}_{k-1}^+ - \sqrt{n + \lambda} \cdot \mathbf{s}_{k-1,i}^+, \quad i = 1, \dots, n \end{aligned} \quad (4.17b)$$

where $\mathbf{s}_{k-1,i}^+$ is the i th column of \mathbf{S}_{k-1}^+ .

3. Propagate the sigma points to obtain:

$$\mathcal{F}_i = \hat{\mathbf{x}}_{k-1}^+ + \int_{t_{k-1}}^{t_k} \mathbf{f}(\hat{\mathbf{x}}(\tau), \tau) d\tau \Big|_{\hat{\mathbf{x}}(t_{k-1})=\mathcal{X}_{i,k-1}^+}, \quad i = 0, \dots, 2n \quad (4.17c)$$

4. Evaluate the a priori state estimate and covariance:

$$\hat{\mathbf{x}}_k^- = \sum_{i=0}^{2n} w_i^m \mathcal{F}_i \quad (4.17d)$$

$$\mathbf{P}_k^- = \sum_{i=0}^{2n} w_i^c (\mathcal{F}_i - \hat{\mathbf{x}}_k^-)(\mathcal{F}_i - \hat{\mathbf{x}}_k^-)^T + \mathbf{Q}_{k-1} \quad (4.17e)$$

where

$$w_0^m = \frac{\lambda}{n + \lambda}, \quad w_0^c = w_0^m + 1 - \alpha^2 + \beta, \quad w_j^m = w_j^c = \frac{1}{2(n + \lambda)}, \quad j = 1, \dots, 2n \quad (4.17f)$$

5. Perform the upper Cholesky decomposition $\mathbf{S}_k^- = \text{chol}(\mathbf{P}_k^-)$, and calculate the sigma points:

$$\begin{aligned} \mathcal{X}_{0,k}^- &= \hat{\mathbf{x}}_k^-, \\ \mathcal{X}_{i,k}^- &= \hat{\mathbf{x}}_k^- + \sqrt{n + \lambda} \cdot \mathbf{s}_{k,i}^-, \\ \mathcal{X}_{i+n,k}^- &= \hat{\mathbf{x}}_k^- - \sqrt{n + \lambda} \cdot \mathbf{s}_{k,i}^-, \quad i = 1, \dots, n \end{aligned} \quad (4.17g)$$

6. Transform the sigma points to obtain:

$$\mathcal{Y}_{i,k}^- = \mathbf{h}(\mathcal{X}_{i,k}^-, t_k), \quad i = 0, \dots, 2n \quad (4.17h)$$

and evaluate the estimated measurement, the estimated measurement covariance, and the estimated state-measurement variance:

$$\hat{\mathbf{y}}_k = \sum_{i=0}^{2n} w_i^m \mathcal{Y}_{i,k}^- \quad (4.17i)$$

$$\mathbf{P}_{yy,k} = \sum_{i=0}^{2n} w_i^c (\mathcal{Y}_{i,k}^- - \hat{\mathbf{y}}_k) (\mathcal{Y}_{i,k}^- - \hat{\mathbf{y}}_k)^T + \mathbf{R}_k \quad (4.17j)$$

$$\mathbf{P}_{xy,k} = \sum_{i=0}^{2n} w_i^c (\mathcal{X}_{i,k}^- - \hat{\mathbf{x}}_k^-) (\mathcal{Y}_{i,k}^- - \hat{\mathbf{y}}_k)^T \quad (4.17k)$$

7. Evaluate the a posteriori state estimate and covariance:

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{y}_k - \hat{\mathbf{y}}_k) \quad (4.17l)$$

$$\mathbf{P}_k^+ = \mathbf{P}_k^- - \mathbf{K}_k \mathbf{P}_{yy,k} \mathbf{K}_k^T \quad (4.17m)$$

where

$$\mathbf{K}_k = \mathbf{P}_{xy,k} \mathbf{P}_{yy,k}^{-1} \quad (4.17n)$$

4.2.3 Cubature Kalman Filter

A more recent reformulation of the Kalman Filter for nonlinear systems is the Cubature Kalman Filter (CKF) proposed by Arasaratnam and Haykin [82]. In their derivation, the authors recall that a state estimate $\mathbf{x} \in \mathbb{R}^n$ and its covariance $\mathbf{P} \in \mathbb{R}^{n \times n}$ are themselves representative of an underlying state probability distribution. While the UKF approximates this probability distribution using a set of sigma points, the CKF approximates the integrals necessary to pass a Gaussian probability distribution through a nonlinear transformation. These multidimensional integrals are approximated by means of a *spherical-radial cubature rule*:

$$\int_{\mathbb{R}^n} \mathbf{g}(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} \approx \sum_{i=1}^m w_i \mathbf{g}(\mathbf{x}_i) \quad (4.18)$$

where $\mathbf{g}(\mathbf{x})$ is our nonlinear transformation, $p(\mathbf{x})$ is the Gaussian probability distribution of our state, w_i are cubature weights, and \mathbf{x}_i are cubature points. In the CKF, the chosen cubature rule results in cubature points defined as:

$$\begin{aligned} \mathbf{x}_i &= \hat{\mathbf{x}} + \sqrt{n} \cdot \mathbf{s}_i, \\ \mathbf{x}_{i+n} &= \hat{\mathbf{x}} - \sqrt{n} \cdot \mathbf{s}_i, \quad i = 1, \dots, n \end{aligned} \quad (4.19)$$

where \mathbf{s}_i is the i th column of $\mathbf{S} = \text{chol}(\mathbf{P})$, the upper Cholesky decomposition of our covariance matrix.

In the equation above, we note that there are a total of $2n$ cubature points. When all points are weighted equally, the cubature weights become:

$$w_i = \frac{1}{2n} \quad i = 1, \dots, n \quad (4.20)$$

With these definitions in hand, we are ready to apply our spherical-radial cubature rule in the context of a nonlinear filter.

The Cubature Kalman Filter

For brevity, we have omitted the CKF algorithm from the current chapter as it is nearly identical to the UKF algorithm. The CKF algorithm can be obtained from the UKF algorithm above by setting $\alpha = 1$, $\kappa = 0$, and $\beta = 0$. These parameters result in uniform sigma point weights $w_j^m = w_j^c = (2n)^{-1}$ for $j = 1, \dots, 2n$ and a

suppressed central sigma point ($w_0^m = w_0^c = 0$).

4.2.4 Continuous-Discrete Kalman Filtering

For our final adaptation of the Kalman filter to nonlinear systems, we will revisit the time update from our UKF summary in Section 4.2.2. Examining (4.17b) through (4.17f), we note that the sigma points are defined at t_{k-1} and propagated forward to t_k individually. It is from these individually propagated sigma points that we obtain our time-updated state estimate $\hat{\mathbf{x}}_k^-$ and covariance \mathbf{P}_k^- .

As an alternative time update, Särkkä [83] proposed the use of the unscented transform to realize continuous dynamics for the physical model in a Kalman filter. Such an approach was found to exhibit numerical stabilities under certain integration constraints³.

To realize this time update, we must redefine our sigma points in continuous time:

$$\begin{aligned}\mathcal{X}_0 &= \hat{\mathbf{x}}, \\ \mathcal{X}_i &= \hat{\mathbf{x}} + \sqrt{n + \lambda} \cdot \mathbf{s}_i, \\ \mathcal{X}_{i+n} &= \hat{\mathbf{x}} - \sqrt{n + \lambda} \cdot \mathbf{s}_i, \quad i = 1, \dots, n\end{aligned}\tag{4.21}$$

where all variables retain their UKF definitions. We can then define a composite function in continuous time which is dependent on these instantaneous sigma points:

$$\hat{\mathbf{f}}(\hat{\mathbf{x}}, t) = w_0^m \mathbf{f}(\mathcal{X}_0, t) + w_1^m \mathbf{f}(\mathcal{X}_1, t) + \dots + w_{2n}^m \mathbf{f}(\mathcal{X}_{2n}, t)\tag{4.22}$$

The function $\hat{\mathbf{f}}(\hat{\mathbf{x}}, t)$ can then be directly propagated to advance our state estimate forward through time. When implemented, such a function will require our numerical integrator to re-evaluate the sigma points at each timestep of integration.

Below, we have outlined the algorithm of Särkkä [ibid.] for a continuous time update. This algorithm propagates $\hat{\mathbf{x}}_{k-1}^+$ and \mathbf{P}_{k-1}^+ into $\hat{\mathbf{x}}_k^-$ and \mathbf{P}_k^- . When this algorithm takes the place of (4.17b) through (4.17f), we achieve a new method of filtering where the time update is achieved in continuous time and the measurement update is achieved in discrete time. This filter is known as the Continuous-Discrete Unscented Kalman Filter (CDUKF).

³Specifically, as the timestep of integration was increased in a fixed-step Runge-Kutta integration scheme, the continuous dynamics outperformed the discrete dynamics [83].

As noted in Section 4.2.3, the algorithm for the Cubature Kalman Filter is nearly identical to that of the UKF. Thus, by replacing the sigma points and weights with cubature points and weights we can realize a Continuous-Discrete Cubature Kalman Filter (CDCKF).

Continuous Time Update for the CDUKF and CDCKF

Note: This algorithm is designed for use with the UKF in Section 4.2.2; all variables retain their definition from that section. To realize the CDCKF, apply the changes detailed in Section 4.2.3.

1. Define the sigma points in continuous time:

$$\begin{aligned}\mathcal{X}_0 &= \hat{\mathbf{x}}, \\ \mathcal{X}_i &= \hat{\mathbf{x}} + \sqrt{n + \lambda} \cdot \mathbf{s}_i, \\ \mathcal{X}_{i+n} &= \hat{\mathbf{x}} - \sqrt{n + \lambda} \cdot \mathbf{s}_i, \quad i = 1, \dots, n\end{aligned}\tag{4.23a}$$

2. Define a continuous function $\hat{\mathbf{f}}$ which is a weighted summation of the physical model $\mathbf{f}(\mathbf{x}, t)$ applied to each sigma point:

$$\hat{\mathbf{f}} = \sum_{i=0}^{2n} w_i^m \mathbf{f}(\mathcal{X}_i, t)\tag{4.23b}$$

3. Define an intermediate matrix $\boldsymbol{\Gamma}$ which will be used in the time derivative of our state covariance:

$$\begin{aligned}\boldsymbol{\Gamma} &= \sum_{i=1}^{2n} w_i^c (\mathcal{X}_i - \hat{\mathbf{x}}) \left(\mathbf{f}(\mathcal{X}_i, t) - \hat{\mathbf{f}} \right)^T \\ &\quad + \sum_{i=1}^{2n} w_i^c \left(\mathbf{f}(\mathcal{X}_i, t) - \hat{\mathbf{f}} \right) (\mathcal{X}_i - \hat{\mathbf{x}})^T \\ &\quad + \mathbf{Q}(t)\end{aligned}\tag{4.23c}$$

where $\mathbf{Q}(t)$ is the continuous process noise covariance.

4. Define a matrix function $\Phi(\mathbf{A})$ which will be used in the time derivative of our state covariance:

$$\Phi(\mathbf{A}) = \text{s-triu}(\mathbf{A}) + \frac{1}{2} \text{diag}(\mathbf{A})\tag{4.23d}$$

where $s_triu(\mathbf{A})$ is the strictly upper triangular part of matrix \mathbf{A} (i.e., the components above the main diagonal) and $diag(\mathbf{A})$ is the diagonal part of \mathbf{A} .

5. With the definitions above, the time derivative of our state estimate $\hat{\mathbf{x}}$ and Cholesky-decomposed state covariance $\mathbf{S} = chol(\mathbf{P})$ can be related as:

$$\frac{d\hat{\mathbf{x}}}{dt} = \hat{\mathbf{f}}, \quad (4.23e)$$

$$\frac{d\mathbf{S}}{dt} = \mathbf{S}\Phi(\mathbf{S}^{-1}\Gamma\mathbf{S}^{-T}). \quad (4.23f)$$

6. Finally, taking all of the previous equations into account we can realize our time update in continuous time:

$$\hat{\mathbf{x}}_k^- = \int_{t_{k-1}}^{t_k} \frac{d\hat{\mathbf{x}}}{d\tau} d\tau \Big|_{\hat{\mathbf{x}}(t_{k-1})=\hat{\mathbf{x}}_{k-1}^+} \quad (4.23g)$$

$$\mathbf{S}_k^- = \int_{t_{k-1}}^{t_k} \frac{d\mathbf{S}}{d\tau} d\tau \Big|_{\hat{\mathbf{x}}(t_{k-1})=\hat{\mathbf{x}}_{k-1}^+, \mathbf{S}(t_{k-1})=\mathbf{S}_{k-1}^+} \quad (4.23h)$$

$$\mathbf{P}_k^- = \mathbf{S}_k \mathbf{S}_k^T \quad (4.23i)$$

4.3 Kalman Flight Heritage

Throughout this chapter, we have implicitly assumed that the Kalman filter and its nonlinear derivatives are suitable for our current application. This assumption has basis in the literature; as Montenbruck [35] notes, Kalman filtering of GPS position fixes is a flight proven technology.

In 1982, Landsat-4 became the first spacecraft to feature an onboard GPS receiver [84]. In the ensuing decades, it has become common for satellite missions in LEO to use GPS signals for the purposes of navigation. However, the analysis of these signals in a Kalman filter to achieve an onboard navigation solution is a more recent development. Notable examples of the latter include:

- the BIRD microsatellite (launched 2001), which demonstrated real-time autonomous navigation by using GPS signals to achieve 5 m (3D RMS) position accuracy in LEO [85];

- the PROBA-2 microsatellite (launched 2009), which employs a commercial-off-the-shelf single-frequency GPS receiver and achieves a real-time position accuracy of 1.1 m (3D RMS) [71]; and
- the relative navigation mission PRISMA (launched 2010), which features two satellites flying the same receiver as PROBA-2 and achieves metre-level absolute positioning and decimetre-level relative positioning [86].

In addition to LEO flight heritage, the Kalman filter has been examined for GPS-based navigation in HEO. These efforts include:

- a 1996 study by Potti et al. [60] for the never-flown Archimedes mission, featuring simple dynamic and measurement models to demonstrate the feasibility of GPS-based navigation in HEO;
- a 2005 hardware-in-the-loop test campaign by Moreau et al. [55] utilizing an RF generator, a physical receiver, and the NASA GPS-Enhanced Onboard Navigation System (GEONS) software to achieve 7 m (3D mean) position accuracy in HEO;
- a 2010 simulation study by Lorga et al. [63] which incorporated advanced tracking algorithms and signals from the Galileo constellation to achieve < 20 m (3D 1σ) accuracy in HEO; and
- a 2010 mathematical study by Li et al. [87] which demonstrated the observability of the problem of GPS-based navigation in HEO when only one pseudorange is available for processing.

It is worth noting that these HEO studies focused solely on the Extended Kalman Filter, firmly establishing the EKF as suitable for this problem. Literature on the other filters is more sparse:

- a study by Choi et al. [88] found only marginal improvement for UKF-based navigation in LEO when compared to EKF-based navigation. Specifically, the 3D RMS position errors for CHAMP improved from 12.9 m to 12.1 m, while those for KOMPSAT-2 improved from 8.7 m to 8.2 m.

- To date, the CKF has not been applied to the problem of GPS-based orbit determination. However, a study by Pesonen and Piché [89] compared the CKF, UKF, and EKF for the problem of Earth-bound positioning using GPS signals fused with signals from mobile phone base stations. In their study, no single filter was found to consistently outperform the others.

Both of these studies demonstrate that frequent measurement updates tend to minimize the benefits of the UKF and CKF over the simpler EKF. However, the problem of satellite navigation in HEO presents long periods of measurement outage over the apogee arc. This will necessitate long period propagation of state and state covariance through the nonlinear equations of motion, a task to which the UKF and CKF may be well-suited. Thus, there exists an opportunity to examine navigation in HEO by other methods of nonlinear filtering. This opportunity will be explored further in Chapter 5.

4.4 Chapter Summary

In this chapter, we have provided an overview of the Kalman filter as an elegant algorithm for state estimation. We have summarized a variety of methods by which the Kalman filter can be applied to a nonlinear system; these methods include the Extended Kalman Filter, the Unscented Kalman Filter, the Cubature Kalman Filter, and Continuous-Discrete Kalman Filters. Finally, we summarized the past applications of Kalman filters for GPS-based satellite navigation, including flight experience in LEO and simulation studies for HEO. With these analytical frameworks established and demonstrated, we are ready to proceed to our final task: evaluation of nonlinear filtering algorithms for the autonomous navigation of satellites in highly elliptical orbit.

Chapter 5

Nonlinear Filter Performance

As discussed in the introduction to this thesis (see Section 1.4), we are completing a navigation study through simulation. The goal of this study is to design a navigation solution that matches a known truth orbit when provided with measurements derived from that orbit. In this chapter, we will assess the accuracy of our navigation solution by comparing it directly to this known truth.

In this chapter, we will employ our propagation models from Chapter 2, our measurement models from Chapter 3, and our filtering algorithms from Chapter 4. These models and algorithms will be fused in a MATLAB-based simulation environment to quantitatively explore the problem of navigation in highly elliptical orbit.

5.1 Simulation Implementation

The bulk of the simulation work was completed in a customized MATLAB environment. The truth orbit was generated using the high-fidelity propagation capabilities of Satellite Tool Kit (STK 9.2.0) from Analytical Graphics Inc. The GPS measurements were generated using a purpose-built GPS simulator, the LabPro-4000 from Navigation Laboratories.

The simulation itself is a six-step process:

1. Simulate a truth orbit using the high fidelity propagator included in AGI STK 9.2.0. The truth orbit is configured with the parameters in Tables 2.8 and 2.9 and initialized with the initial conditions in Table 2.3.
2. Using MATLAB, convert the truth orbit into a position and attitude trajectory readable by the GPS simulator.

3. Initialize the GPS simulator using the formatted truth orbit trajectory.
4. Perform a GPS simulation¹ and save the resulting measurement data products.
5. Convert the data products into pseudorange measurements by calculating SV visibility and the link budget in MATLAB.
6. Provide the measurements to a MATLAB-based navigation simulation to test a given filter configuration.
7. Compare the navigation solution to the original STK truth orbit.

The realism of this simulation is enhanced by the use of separate dynamic models for the truth and navigation solution. We treat the propagator discrepancies as process noise stemming from differences in physical model fidelity. Great care was taken to eliminate other discrepancies, including those between physical constants and reference frames.

5.2 Preamble: Setting up the Simulation

During the simulation setup, a number of design decisions were implemented. Some were based on the literature while others were bound by the scope of the current study. These decisions are summarized in Table 5.1 and discussed further in the following three sections. With these in hand, we are ready to begin the examination of our navigation solution.

5.2.1 Simulation Configuration

Simulation length: The simulation focuses on one twelve-hour Molniya orbit. Due to excellent GPS coverage near the Earth, the navigation solution converges to metre-level accuracy at each perigee regardless of the inaccuracies accumulated over the previous apogee arc. Thus, we only need to examine the solution's behaviour as it traverses the northern arc of our Molniya orbit.

Truth orbit: For our truth model we have configured STK to provide its maximum level of realism. This includes all primary force models from Table 2.8 and all secondary force models discussed in Section 2.2.5.

¹For details on the simulator operation, see Appendix C.1.

Filter Algorithm: As discussed in Chapter 4, our baseline algorithm is the Extended Kalman Filter (EKF). The other methods of nonlinear filtering introduced in that chapter are discussed in Section 5.5.

5.2.2 Propagator Configuration

Physical model: In Chapter 2 we developed a number of different force models for our propagator. In our baseline case, we include only the two most significant forces: geoidal Earth gravity (EGM-96 10×10) and third-body gravity. Our physical model $\mathbf{f}(\mathbf{x}, t)$ in (4.6a) therefore takes the form:

$$\mathbf{f}(\mathbf{x}, t) = \mathbf{a}_{\text{geoid}}(\mathbf{x}, t) + \mathbf{a}_{3,\odot}(\mathbf{x}, t) + \mathbf{a}_{3,\mathbb{C}}(\mathbf{x}, t) \quad (5.1)$$

where all variables are defined in Table 2.6 and the geoidal acceleration $\mathbf{a}_{\text{geoid}}(\mathbf{x}, t)$ is calculated using the EGM-96 10×10 coefficients on the WGS-84 ellipsoid. The impact of propagator fidelity will be further examined in Section 5.4.2.

Initial Conditions: To demonstrate filter convergence, the initial conditions for our state vector \mathbf{x}_0^+ were randomly generated using MATLAB's `randn` function. The standard deviations used for this random initialization can be found in Table 5.1. The initial state covariance \mathbf{P}_0^+ is initialized with the corresponding variances along the main diagonal.

Process noise: A necessary prerequisite for EKF-based state estimation is the specification of the process noise covariance \mathbf{Q}_{k-1} implemented in the filter. This value approximately characterizes the actual process noise which results from the differences between our truth and filter propagators. An iterative method of filter tuning resulted in the following discrete-time² process noise covariance matrix:

$$\mathbf{Q}_k = \begin{bmatrix} Q_r \cdot \mathbf{I}_3 & 0 & 0 \\ 0 & Q_v \cdot \mathbf{I}_3 & 0 \\ 0 & 0 & \mathbf{Q}_{\text{clock}} \end{bmatrix} \quad (5.2)$$

²The continuous-time filters discussed in Section 5.5 will use the continuous equivalent of (5.2).

where Q_r is our position process noise, Q_v is our velocity process noise, \mathbf{Q}_{clock} is the receiver clock process noise, and \mathbf{I}_3 is a 3×3 identity matrix. The values for Q_r and Q_v are constant³ and can be found in Table 5.1. The definition of \mathbf{Q}_{clock} will depend on the current mode of receiver clock operation:

- If the clock is drifting, the discrete-time equivalent of (3.17) will be used.
- If the clock is being steered, $\mathbf{Q}_{clock} = [\sigma_{steer}^2]$ using the filter value for σ_{steer} as given in Table 3.1.

Forces in $\mathbf{F}(\mathbf{x}, t)$: We recall from (4.7a) that our required Jacobian matrix $\mathbf{F}(\mathbf{x}, t)$ is reliant on the physical model $\mathbf{f}(\mathbf{x}, t)$. We should therefore have a term in $\mathbf{F}(\mathbf{x}, t)$ corresponding to each term in (5.1). However, for this orbital regime a simplified Jacobian can be implemented without loss of filter accuracy. This Jacobian includes only two factors: the two-body acceleration (2.2) and the current mode of receiver clock operation:

$$\mathbf{F}(\mathbf{x}, t) = \begin{bmatrix} 0 & \mathbf{I}_3 & 0 \\ \frac{3\mu_{\oplus}}{r^5} \mathbf{r} \mathbf{r}^T & 0 & 0 \\ 0 & 0 & \mathbf{F}_{clock} \end{bmatrix} \quad (5.3)$$

where μ_{\oplus} is the gravitational parameter of the Earth from Table 2.7, \mathbf{r} is the position vector of our satellite, and $\mathbf{F}_{clock} = \mathbf{F}_{steer}$ or \mathbf{F}_{drift} as defined in (5.4). A more complex Jacobian accounting for the Earth oblateness⁴ was examined through simulation but not implemented as it was found to have no measurable impact on the solution.

The receiver clock Jacobian will depend on the current mode of receiver clock operation. In accordance with (3.15) and (3.16), we derive:

$$\mathbf{F}_{steer}(\mathbf{x}, t) = \begin{bmatrix} -\tau_s^{-1} \end{bmatrix} \quad \mathbf{F}_{drift}(\mathbf{x}, t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad (5.4)$$

where τ_s is the coloured-noise decay constant as defined in Table 3.1.

³Note that these values are specific to our 30 s measurement interval.

⁴More commonly referred to as the J_2 effect, this was achieved by modelling a 2×0 EGM-96 gravity field.

5.2.3 Measurement Configuration

Measurement model: In Chapter 3 we presented truth and filter models for the GPS pseudorange measurement. Our measurement model $\mathbf{h}(\mathbf{x}, t)$ in (4.6b) therefore takes the form:

$$\mathbf{h}(\mathbf{x}, t) = \mathbf{h}_{GPS}^\Delta(\mathbf{x}, t) + \mathbf{h}_{Rclk}(\mathbf{x}, t) \quad (5.5)$$

where all variables are defined in Table 3.4.

Measurement model Jacobian: We recall from (4.7b) that our required Jacobian matrix $\mathbf{H}(\mathbf{x}, t)$ is reliant on the measurement model $\mathbf{h}(\mathbf{x}, t)$. Based on our definition in (5.5), we define:

$$\mathbf{H}(\mathbf{x}, t) = \text{col}_i \left[\begin{pmatrix} \mathbf{r} - \mathbf{r}_{SV}^\Delta \\ \frac{\partial}{\partial \mathbf{r}} \end{pmatrix}^T \quad 0 \quad 0 \quad \mathbf{H}_{clock} \right] \quad (5.6)$$

where \mathbf{r} is the position vector of our satellite, \mathbf{r}_{SV}^Δ are the perturbed SV ephemerides, and $\mathbf{H}_{clock} = \mathbf{H}_{steer}$ or \mathbf{H}_{drift} as defined in (5.7). The latter equation is given by:

$$\mathbf{H}_{steer}(\mathbf{x}, t) = [c] \quad \mathbf{H}_{drift}(\mathbf{x}, t) = [c \quad 0] \quad (5.7)$$

where c is once again the speed of light.

Measurement frequency: Current spaceborne GPS receivers can provide raw pseudoranges at a frequency of 1 Hz – see, for example, the documented on-orbit performance of the Astrium MosaicGNSS receiver [90]. For this study, a conservative update interval of 30 s was selected to ensure the robustness of the navigation solution. Such an interval was found to be sufficient for LEO by Montenbruck et al. [35] in their analysis for filter-based navigation of PROBA-2.

GPS receiver channels: GPS receivers have a hardware-imposed limit on the number of GPS satellites they can track at a given time. We again used the example of the Astrium MoisaicGNSS receiver [90] and include a maximum of eight pseudoranges in our navigation solution.

GPS receiver antenna field-of-view: In this study, a nadir-pointing attitude was assumed over our entire HEO orbit. This allowed us to consider a nadir-pointing

antenna and a zenith-pointing antenna for direct comparison with the results of Potti et al. [60]. Both antennae were assumed to have a full hemisphere of visibility (2π steradians) with a gain pattern given by Figure 3.2. The impact of antenna position is examined in Section 5.4.1.

Measurement noise: As with the process noise, we must provide our filter with the measurement noise covariance \mathbf{R}_k which approximately characterizes the difference between our truth and filter measurement models. As discussed throughout Chapter 3, we expect our GPS pseudorange measurements to be accurate to within tens of metres. An iterative method of filter tuning resulted in the following measurement noise covariance matrix:

$$\mathbf{R}_k = \left[R_{GPS} \cdot \mathbf{I}_k \right] \quad (5.8)$$

where $R_{GPS} = (10 \text{ m})^2 + R_{iono}$ contains the tuned parameter plus an estimate of the ionospheric measurement noise⁵ and \mathbf{I}_k is an identity matrix with dimensions matching the number of currently locked satellites (see Section 3.6).

5.3 Baseline EKF Analysis

Following from the discussion in the previous section, a baseline set of simulations was conducted to verify the navigation solution. Parameters for these simulations were set in accordance with Table 5.1 according to the rationale given throughout Section 5.2. The results of a sample simulation can be seen in Figure 5.1 to Figure 5.6 inclusive.

5.3.1 What do these figures mean?

Figure 5.1 to Figure 5.4 contain the time histories of the state errors and state uncertainties. The former is represented by the solid lines, which trace out the difference ($\hat{\mathbf{x}}_k - \mathbf{x}_k$). The latter is represented by the dashed lines, which define the 3σ bounds.

⁵In line with the discussion in Section 3.4.2, R_{iono} is the square of the i th component of $(1 - \overline{IF}) \cdot \mathbf{h}_{iono}^t(\mathbf{x}, t)$ where \overline{IF} is the average ionosphere factor of 0.6. In effect, the filter assumes that the ionospheric model removes 60% of the ionospheric effect, when in reality it removes between 40–80%.

Table 5.1: Baseline Simulation Configuration

Simulation Settings	
simulation length	12 h
truth model	full STK force model
filter algorithm*	EKF
Propagator Settings	
propagator force models*	EGM-96 (10×10) luni-solar
Jacobian matrix $\mathbf{F}(\mathbf{x}, t)$	two-body (5.3)
process noise covariance \mathbf{Q}_k	$Q_r = 1 \times 10^{-7} \text{ km}^2$ $Q_v = 1 \times 10^{-13} \text{ km}^2 \cdot \text{s}^{-2}$ \mathbf{Q}_{clock} per (5.2)
initial conditions $\mathbf{x}_0^+, \mathbf{P}_0^+$	$r_0 = N(0, 100 \text{ m})$ $v_0 = N(0, 1 \text{ m} \cdot \text{s}^{-1})$ $b_0 = N(0, 1 \mu\text{s})$ $d_0 = N(0, 1 \mu\text{s} \cdot \text{s}^{-1})$
Measurement Settings	
measurement interval	30 s
measurement noise covariance \mathbf{R}_k	$R_{GPS} = (10 \text{ m})^2$
number of GPS antennae*	2 (nadir & zenith)
number of GPS channels	8

*considered for further analysis.

These bounds rely on the definition of our standard deviation vector:

$$\sigma_k = \text{col}_i [\sqrt{\mathbf{p}_{ii,k}}] \quad (5.9)$$

where $\mathbf{p}_{ii,k}$ is the i th diagonal component of \mathbf{P}_k and col_i denotes a column vector where the i th row is given inside the brackets. The dashed lines therefore represent our positive and negative 3σ bounds within which our solution should remain the majority of the time.

Figure 5.5 is a layered figure that combines position error data from Figure 5.1

with satellite visibility data from Figure 3.3. The vertical bars indicate the number of GPS SVs being processed by the navigation solution. A white bar corresponds to zero SVs, a black bar indicates four or more, and the varying shades of grey cover the intermediate values.

Finally, Figure 5.6 contains a quantification of filter convergence. In this figure, we have plotted the 2-norm of the residuals, the difference $(\mathbf{y}_k - \mathbf{h}(\hat{\mathbf{x}}_k^-, t_k))$ used in the Kalman update. This parameter provides insight to the degree of accuracy with which the filter is estimating the measurements it is receiving.

Note on the choice of reference frame: The position and velocity errors are presented in the along/across/radial reference frame. This frame is well-suited to highlight the impact of changes to the navigation solution and is the frame-of-reference for the STDP accuracy requirements (see Section 1.3.3). Following the notation of Vallado [26], this frame is formally defined as:

$$\hat{\mathbf{r}} = \frac{\mathbf{r}}{|\mathbf{r}|} \quad (5.10)$$

$$\hat{\mathbf{s}} = \hat{\mathbf{w}} \times \hat{\mathbf{r}} \quad (5.11)$$

$$\hat{\mathbf{w}} = \frac{\mathbf{r} \times \mathbf{v}}{|\mathbf{r}||\mathbf{v}|} \quad (5.12)$$

where we define:

$\hat{\mathbf{r}}$ radial unit vector, unitless

$\hat{\mathbf{s}}$ along-track unit vector, unitless

$\hat{\mathbf{w}}$ across-track unit vector, unitless

\mathbf{r} instantaneous ECI position⁶, km

\mathbf{v} instantaneous ECI velocity, km·s⁻¹

Note that by these definitions, $\hat{\mathbf{w}}$ is normal to the orbit plane and $\hat{\mathbf{s}}$ points in the same general direction as the velocity vector (they are co-aligned only under certain circumstances). Using these definitions, we can define a rotation matrix to transform

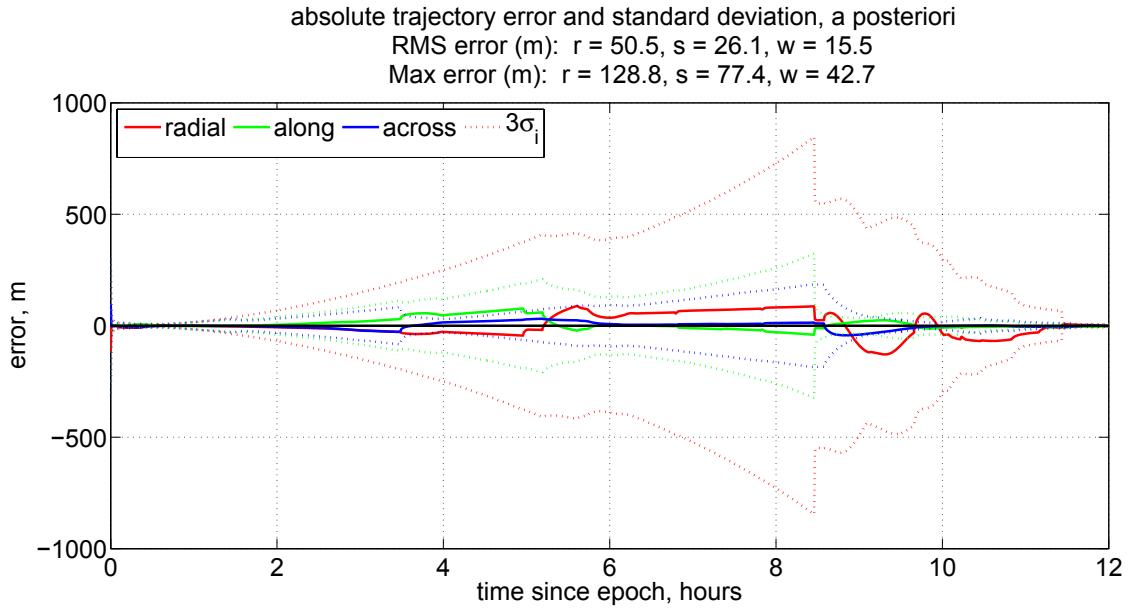


Figure 5.1: Baseline simulation, position error, along/across/radial frame*.

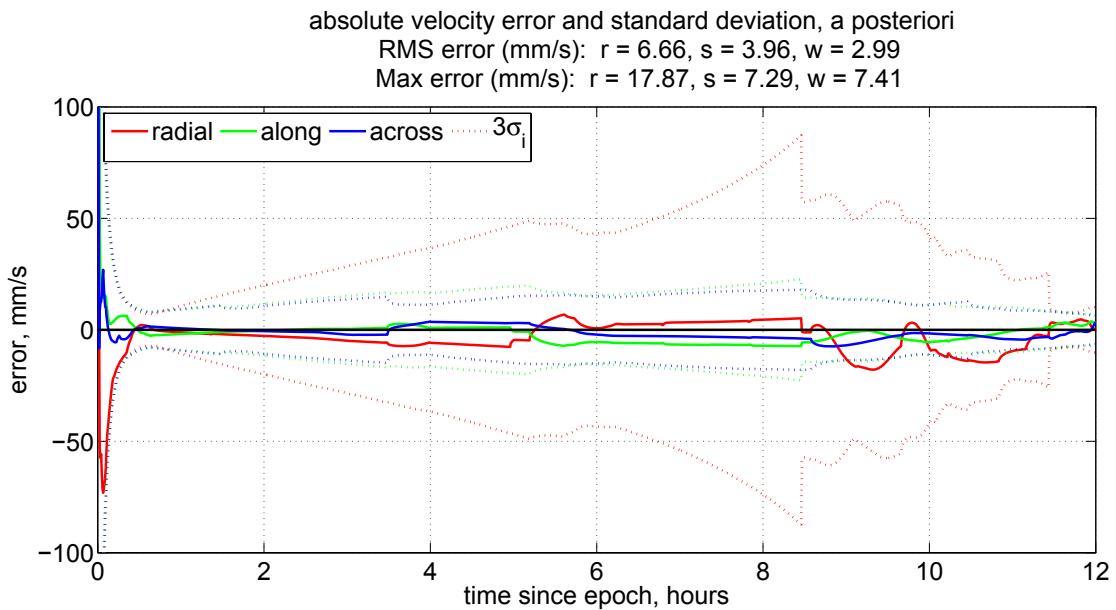


Figure 5.2: Baseline simulation, velocity error, along/across/radial frame*.

(*for a discussion on the along/across/radial reference frame, see Note above.)

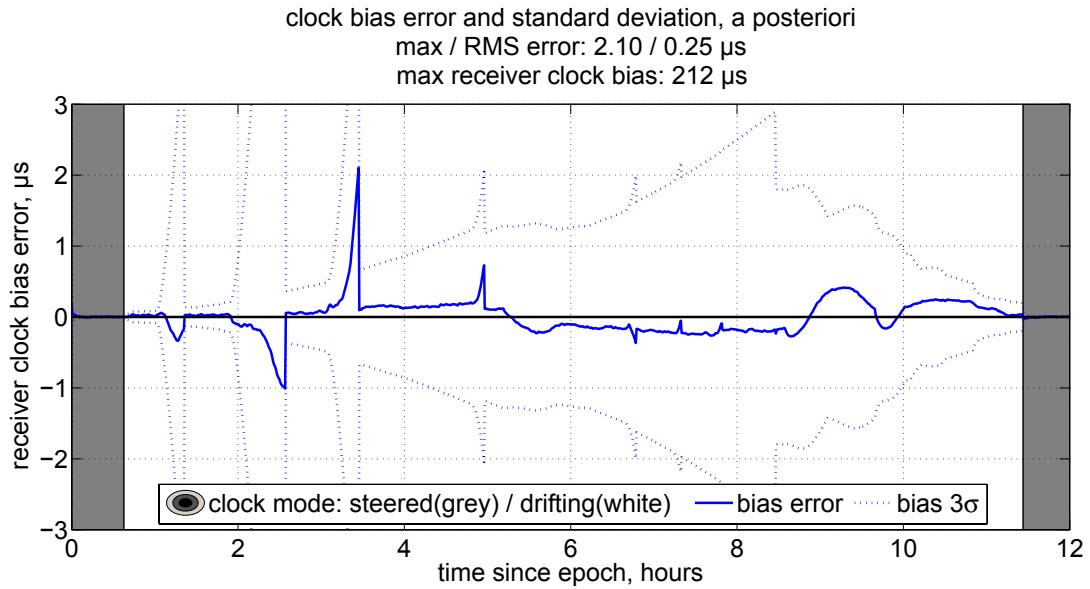


Figure 5.3: Baseline simulation, receiver clock bias error and receiver clock mode.

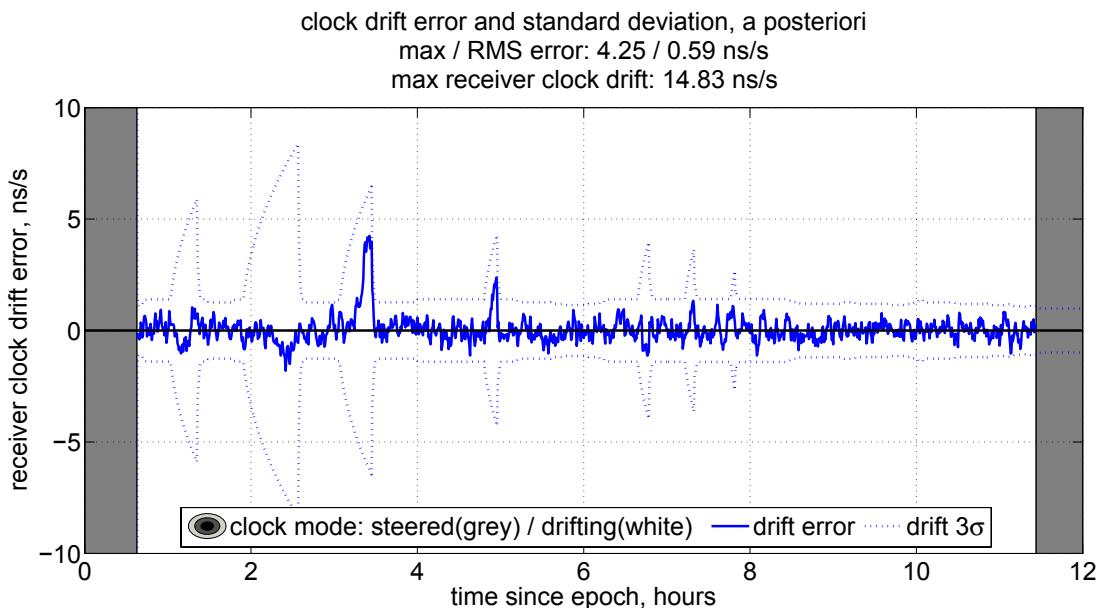


Figure 5.4: Baseline simulation, receiver clock drift error and receiver clock mode.

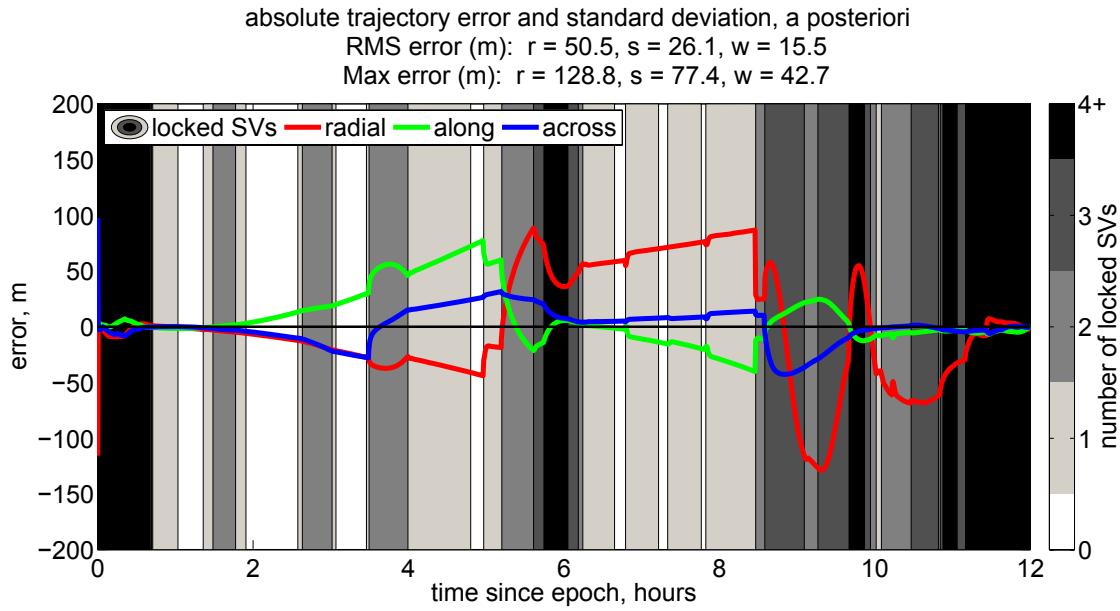


Figure 5.5: Baseline simulation, position error and locked GPS satellites.

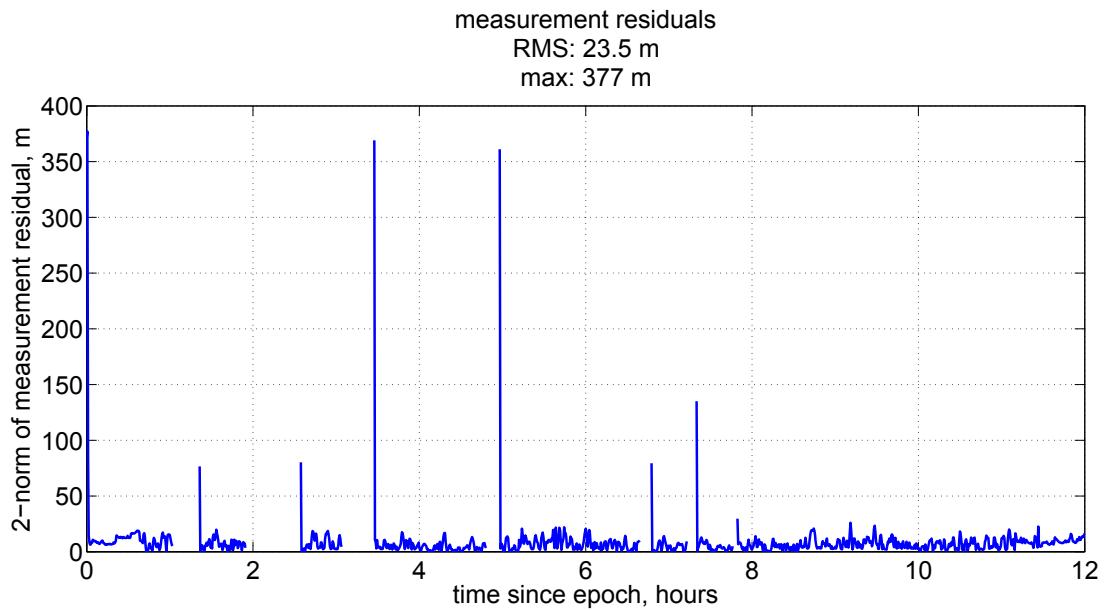


Figure 5.6: Baseline simulation, measurement residuals.

our position and velocity errors and their associated covariance matrices into the desired frame.

5.3.2 Baseline EKF Performance

Examining the results of our baseline simulation in Figures 5.1 to 5.6, we can see that the EKF is working as expected; the solution error is constrained by the 3σ bounds and remains converged throughout the period of simulation. Beyond this general conclusion, each plot contains specific details about our navigation solution.

Position and velocity: Bearing in mind the time period of interest for our simulation (twelve hours from perigee to perigee), we note in Figure 5.1 and Figure 5.2 a steady increase in uncertainty as the satellite traverses the apogee arc of our HEO orbit. This is due to the “dead reckoning” of our solution through this arc of low GPS coverage. The gradual return of coverage from the eighth hour onward is reflected by a corresponding reduction in solution uncertainty, with position accuracy fully converged by the time perigee is reached. As noted by the above-plot statistics, the solution is accurate to tens of metres in position and millimetres per second in velocity.

Receiver clock bias and drift: Figure 5.3 and Figure 5.4 highlight the performance of our receiver clock error as a function of time. Both plots include a grey/white colour scheme to indicate when the receiver clock is drifting (white) and when it is being steered (grey).

Looking at the bias, we note that it converges at perigee with significant deviation only occurring during GPS signal outage. This fact is mirrored in the drift plot, where we see drift error and uncertainty sharply increase during GPS signal outage. It is worth noting that the solution does not estimate the drift around perigee - as detailed in Section 3.5, drift is only estimated during times when the GPS clock is not being steered.

In Figure 5.3, we further note that our receiver clock bias has a maximum error of 2100 ns and an RMS error of 250 ns; the latter translates to a large pseudorange error of 75 m. However, these bias error statistics are computed using data from the

⁶Earth Centred Inertial reference frame; see Appendix A.2 for a discussion on reference frames and coordinate transformations.

entire orbit. A more careful calculation of these statistics ignores the clock bias error during total GPS outage as this freely-drifting bias has no impact on the navigation solution. If these periods of total outage are removed from the statistic calculation, we obtain a maximum clock bias of 410 ns and an RMS value of 160 ns, translating to 48 m of pseudorange inaccuracy.

Further analysis: The remaining two figures offer additional insight on the performance of our navigation solution. In Figure 5.5, we can see that the sharp changes in our navigation solution correspond directly to the acquisition of an additional GPS signal. As each GPS signal has a unique set of ephemeris and clock errors, the minimum variance state estimate will change significantly with the acquisition or loss of a GPS satellite.

The final baseline plot, Figure 5.6, demonstrates that upon signal acquisition, the EKF quickly forces the residual down to the magnitude of the measurement error itself. This confirms the EKF as well-suited for this application.

5.4 Perturbing the Baseline EKF

Having established a nominal performance for our EKF-based navigation solution, we will examine the impact of various options available to optimize the performance of our filter. We will also examine the inclusion of an additional measurement type, passive ranging, to determine its ability to improve our navigation solution.

5.4.1 Effect of Receiver Antenna Positioning

As noted in Section 5.2.3, this study assumes a nadir-pointing spacecraft attitude. In our SV visibility analysis (see Section 3.6.2) we considered two distinct GPS receiver antenna positions on the spacecraft: a nadir-pointing antenna and a zenith-pointing antenna. The baseline case included both antennas. To examine the impact of antenna failure, a simulation was completed with each antenna acting alone. The results can be seen in Figure 5.7 and Figure 5.8.

The zenith-pointing antenna performs as expected; with no signals received during the apogee arc, the solution merely propagates until GPS signals are restored near perigee. More interesting is the nadir-pointing case, where GPS coverage at perigee is sufficient for an accurate navigation solution (see Note below). However, gaps early

in the orbit cause the solution to diverge more rapidly than the two-antenna case. We are therefore comfortable confirming the need for two GPS antennae on our HEO spacecraft.

An additional benefit of this analysis is the demonstration of propagator accuracy in the absence of measurements through the apogee arc. This validates the work done in Chapter 2 and highlights the filter’s ability to reconverge even after multiple hours without a measurement update.

Note on nadir SV visibility: It is at first surprising to see such good GPS signal reception by our nadir-pointing antenna at perigee. However, geometrical calculations show that the Earth-blockage at our perigee altitude of 500 km is only 66.4%, allowing a full third of GPS signals to reach our nadir-pointing antenna. This blockage decreases to 50% at approximately 1100 km altitude (occurring 8 min after perigee) and 22.4% when our spacecraft leaves GPS main lobe coverage at approximately 3900 km altitude (occurring 21 min after perigee). Thus, even for the nadir-pointing antenna, Earth-blockage rapidly becomes a minor source of GPS signal blockage.

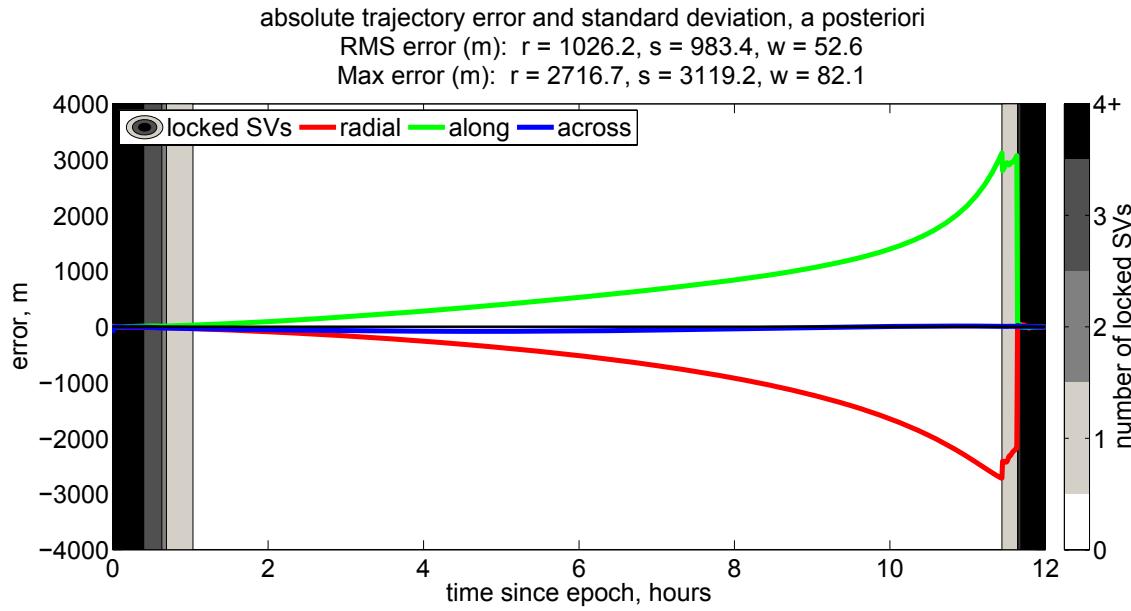


Figure 5.7: Zenith-pointing antenna, position error and locked GPS satellites.

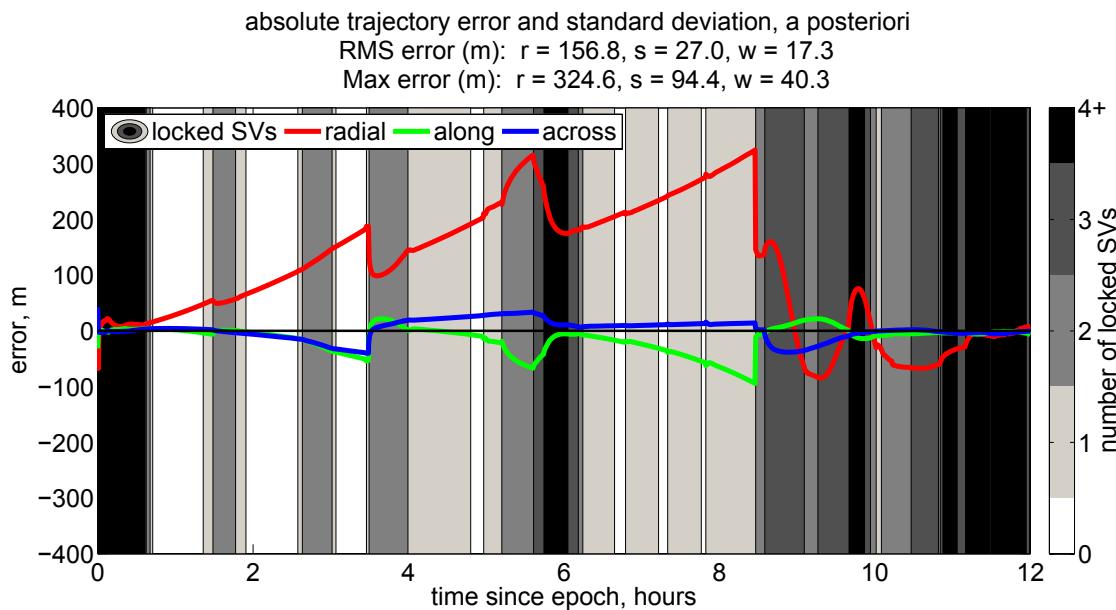


Figure 5.8: Nadir-pointing antenna, position error and locked GPS satellites.

5.4.2 Effect of EKF Propagator Fidelity

Computational efficiency is a necessity for autonomous navigation solutions acting in real-time. With this motivation, two additional propagator configurations were examined:

- a **simple** propagator, consisting only of the oblate Earth gravity model (i.e. the J_2 effect) and the luni-solar perturbation;
- a **complex** propagator, consisting of a 10×10 EGM-96 gravity model, luni-solar perturbations, atmospheric drag, and solar radiation pressure.

These propagators were simulated under identical initial conditions and compared to the baseline case documented in Figure 5.1.

The results for the simple propagator can be seen in Figure 5.9. With a filter runtime of 40.3 s (needed for processing measurements and propagating the solution over the entire orbit) it nearly halves the 79.8 s performance of the baseline case for our 12 h orbit. However, the radial and along-track errors are noticeably larger.

In contrast, the complex propagator as documented in Figure 5.10 improves our navigation error slightly with a corresponding slight increase in computational cost (83.8 s). It is worth noting that this solution is reliant on user-defined coefficients for drag and solar radiation pressure (see Table 2.7); these coefficients will vary from application to application.

We can therefore confirm the suitability of the baseline propagator configuration as an appropriate balance between accuracy and computational load. Specifically, we have demonstrated that accurate navigation solutions can be achieved despite the lack of models for atmospheric drag and solar radiation pressure. This design decision matches that implemented by Berthias et al. [91] for their DIOGENE algorithm, an EKF-based navigation solution which later flew on a number of satellites including PROBA-2 [35].

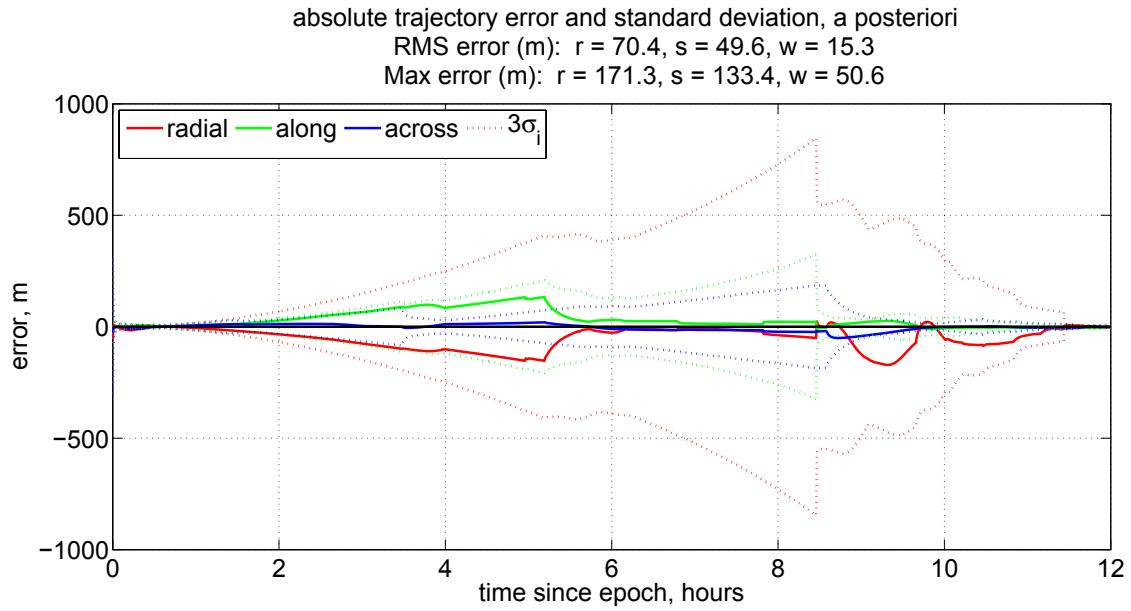


Figure 5.9: Simple propagator, position error.

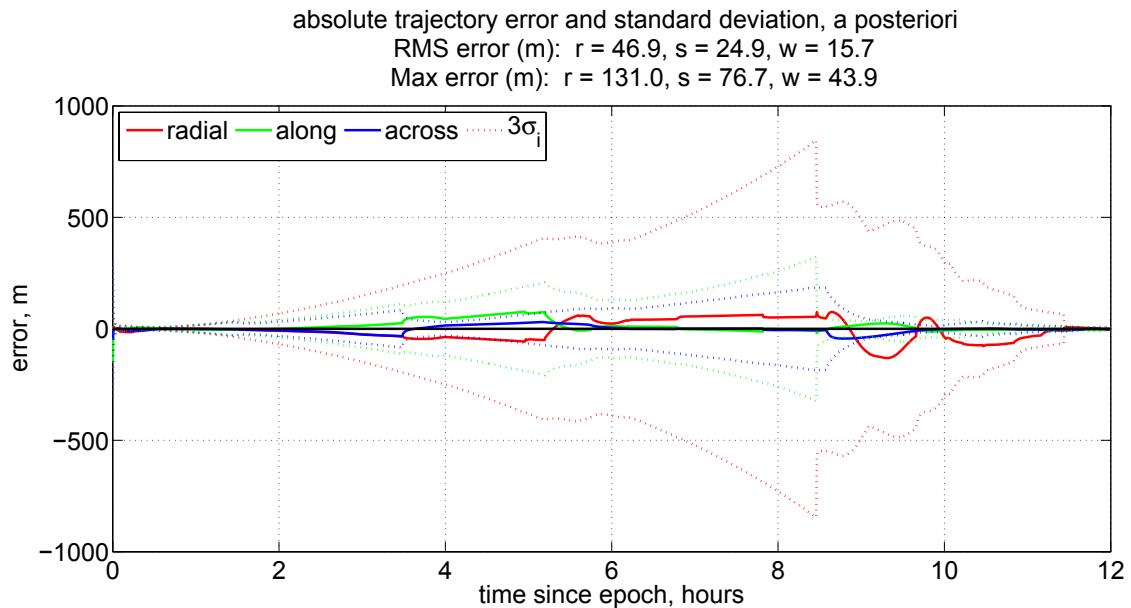


Figure 5.10: Complex propagator, position error.

5.4.3 Effect of Sidelobes

In Section 3.6.3, we introduced the concept of GPS antenna sidelobes. It is anticipated that the inclusion of sidelobes in the GPS transmitter pattern will increase the number of measurements available for navigation. We therefore completed two runs from identical initial conditions; Figure 5.11 uses the gain pattern from Figure 3.2 (sidelobes not included), while Figure 5.12 uses the pattern from Figure 3.6 (sidelobes included).

Comparing the two figures, we notice that the inclusion of sidelobes increases the number of GPS signals available for navigation during the first part of the orbit. This delays the divergence of the navigation solution, resulting in the improvement of the 3D RMS position error from 55.6 m to 43.6 m.

However, it is worth noting that the inclusion of sidelobes does not appear to have an impact on GPS visibility during the high apogee arc. This is due to the fact that our nadir-pointing antenna is focused squarely on the centre of the Earth - a non-ideal configuration for the reception of sidelobe signals. An attitude configuration more friendly to these signals would be required to better assess their ability to impact the navigation solution.

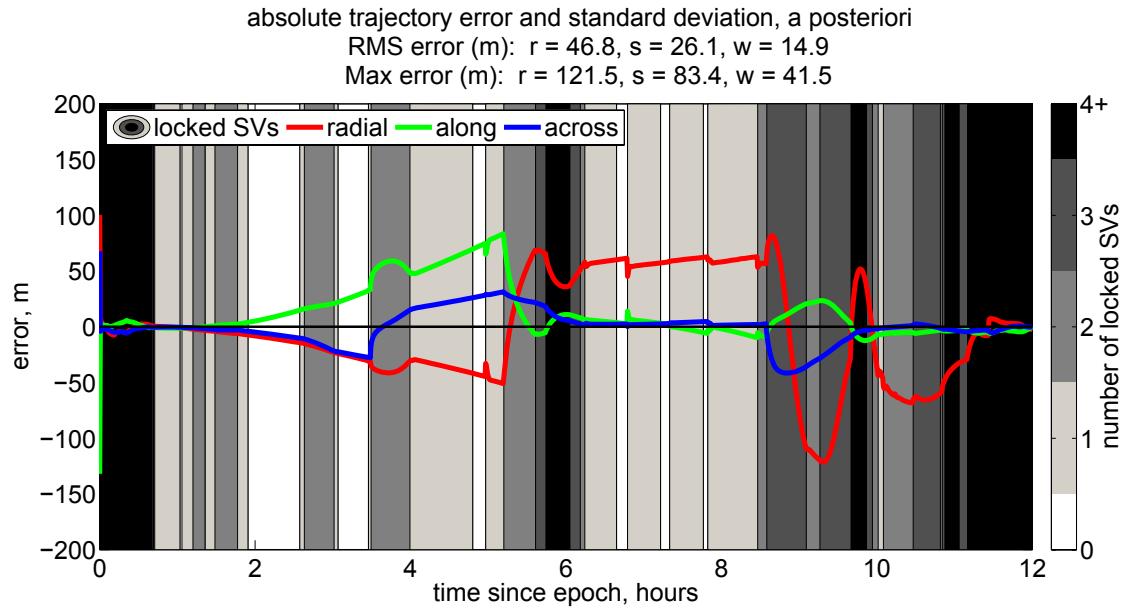


Figure 5.11: Sidelobes OFF, position error and locked GPS satellites.

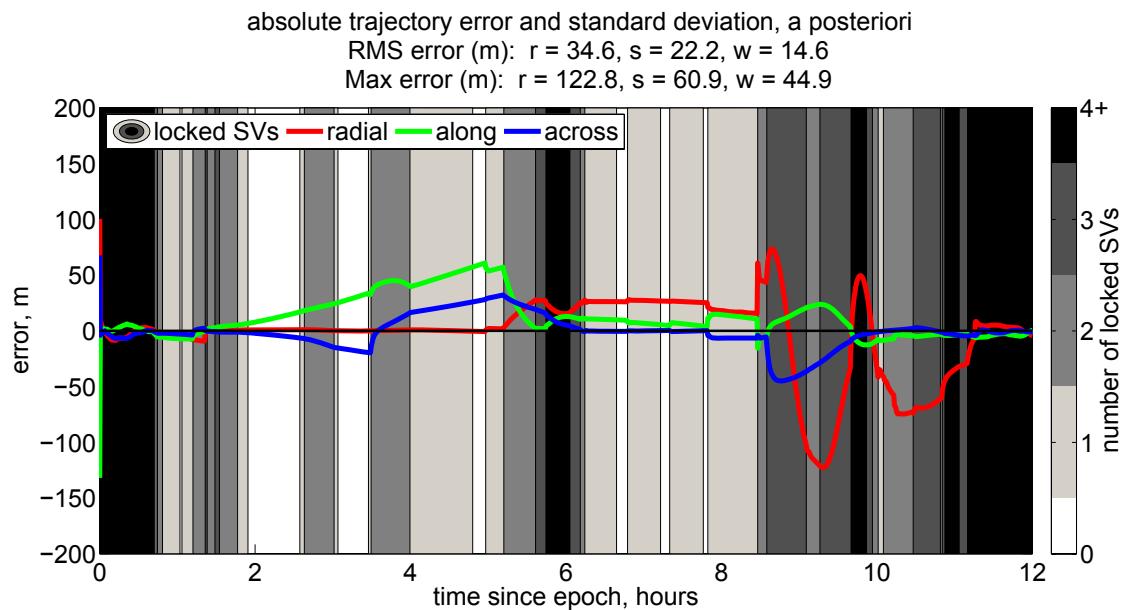


Figure 5.12: Sidelobes ON, position error and locked GPS satellites.

5.4.4 Effect of Receiver Clock Error

To further explore the behaviour of this navigation solution, a simulation was performed where the receiver clock was assumed error-free. This was done to highlight the effect of receiver clock error on the navigation solution. The results of this simulation can be seen in Figure 5.13 and Figure 5.14.

We notice a remarkable improvement in this simulation. RMS position errors drop to less than twenty metres, while no coordinate reaches an error exceeding sixty metres. The solution is well-behaved; error grows during periods of propagation and shrinks (sometimes with great speed) as new measurements are processed. We again observe the transient divergence of the solution when new SVs are acquired (e.g., just before the 8 h mark) – as before, this results from the incorporation of uniquely biased measurements into the navigation solution.

Most interesting is the behaviour of the radial coordinate. Previously the most error-prone, it has an RMS error of a mere 4.4 m – a full order of magnitude improvement! This behaviour sheds light on the behaviour exhibited by radial position error during the baseline simulation. When the satellite is high on the apogee arc, it is only receiving signals from GPS SVs on the far side of the Earth. This line-of-sight is very nearly aligned with the Earth-centred radial direction from the along-across-radial reference frame. Thus, pseudoranges on the apogee arc are primarily functions of two factors: the radial distance plus the (scaled) receiver clock bias.

If the filter is underestimating one of these parameters, it will respond by overestimating the other one. A close examination of Figure 5.1 and Figure 5.3 between the eighth and tenth hour reveals that receiver clock bias and radial position error are indeed negatively correlated.

Thus, we are finally able to reconcile the residuals in Figure 5.6 (where residuals were seen to be well below the twenty metre threshold) and the hundred-plus metre error in radial position between the eighth and tenth hour. The explanation lies in the receiver clock error and its maximal error of 410 ns (approx. 120 m). This behaviour will require further analysis.

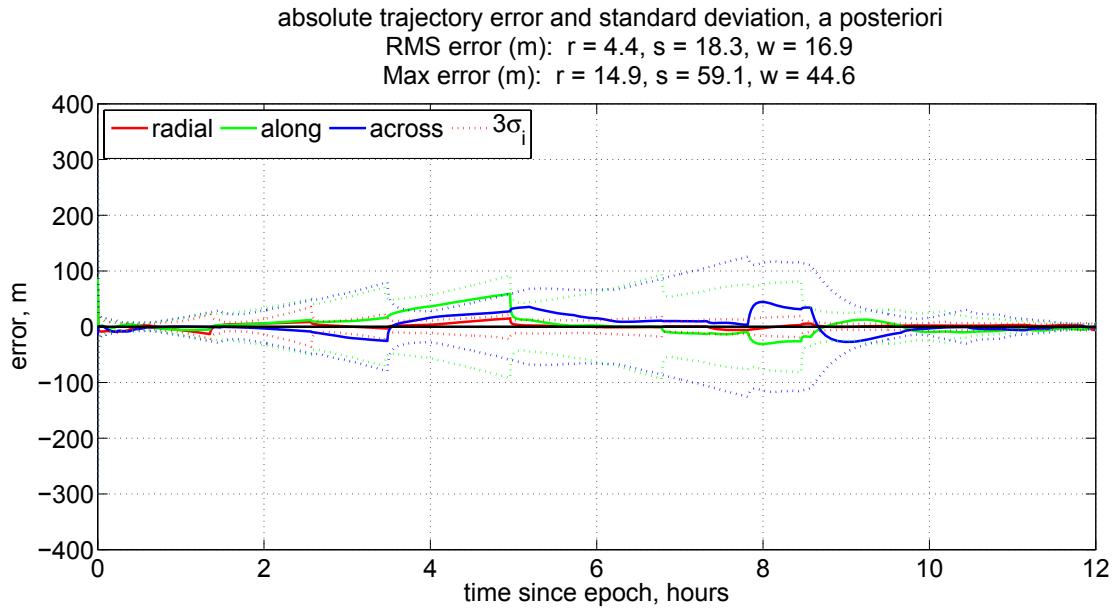


Figure 5.13: Receiver Clock Perfect, position error.

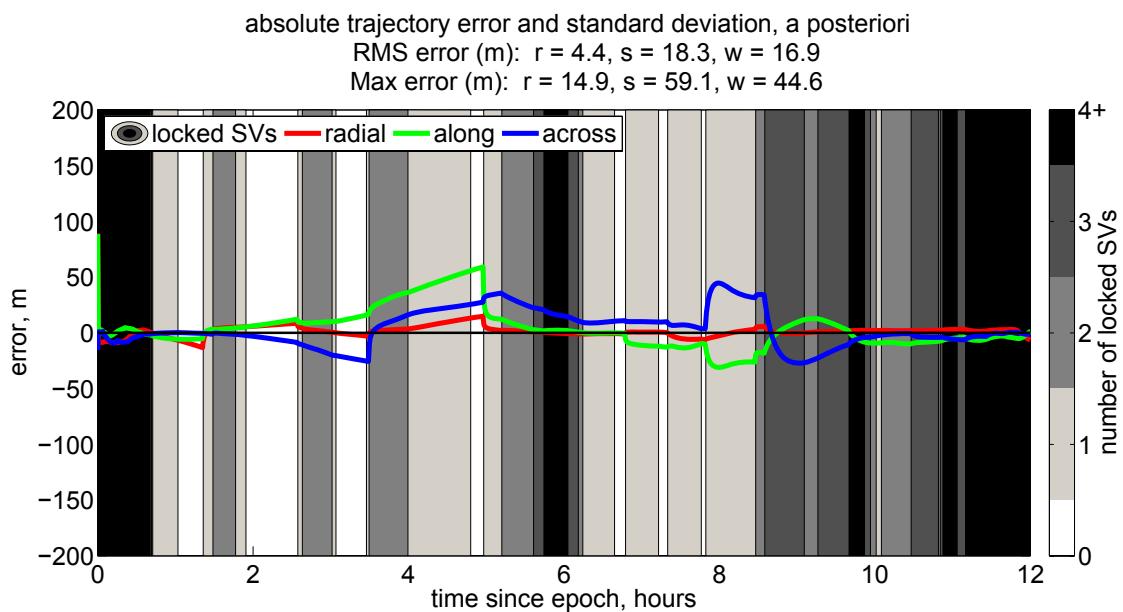


Figure 5.14: Receiver Clock Perfect, position error and locked GPS satellites.

5.4.5 Inclusion of Passive Ranging

In Section 3.7, we introduced an implementation of ranging suited to our goal of autonomous navigation. This passive ranging scenario assumes that two beacons have been installed at existing Canadian Space Agency ground stations. These beacons are specified to broadcast a signal appropriate for pseudorange measurements using the same operating characteristics as a GPS SV transmitter with sidelobes included (see Figure 3.6). To generate measurements for our simulation, we will employ the geometric pseudorange per (3.2) corrupted with a zero-mean white noise signal with a standard deviation of $\sigma_{rge} = 5 \text{ m}$. This value corresponds to the approximate level of pseudorange error that remains after the errors listed in Section 3.7 have been accounted for (cf. United States Coast Guard [92], Table 3-1). For consistent fusion with the existing GPS measurements, we will adopt the same measurement interval of 30 s.

Testing the range measurements: In the previous section, we demonstrated that inaccuracies in the receiver clock correction are responsible for a significant portion of the position error in our navigation solution. Thus, to test the ranging system we ignored receiver clock error and set $R_{rge} = (5 \text{ m})^2$, matching the white noise with which the range measurements are corrupted. The results of that simulation can be seen in Figure 5.15.

Comparing this simulation with Figure 5.14, our previous receiver-clock-perfect simulation, we notice a significant improvement in all three components of position error. The continuous measurement coverage softens the transients that occur when a new GPS signal is acquired. The solution, however, is slightly less smooth; this is expected as our ranging measurements are corrupted by a high-frequency white noise while our GPS measurements are corrupted by the slowly-varying (and more realistic) biases detailed in Section 3.4.

Implementing the range measurements: Having verified the impact of our range measurements in an idealized scenario, we included receiver clock error per (3.12) and retained our measurement noise of $R_{rge} = (5 \text{ m})^2$. The results of this simulation can be seen in Figure 5.16.

The results of this simulation are somewhat counter-intuitive. Comparing it to

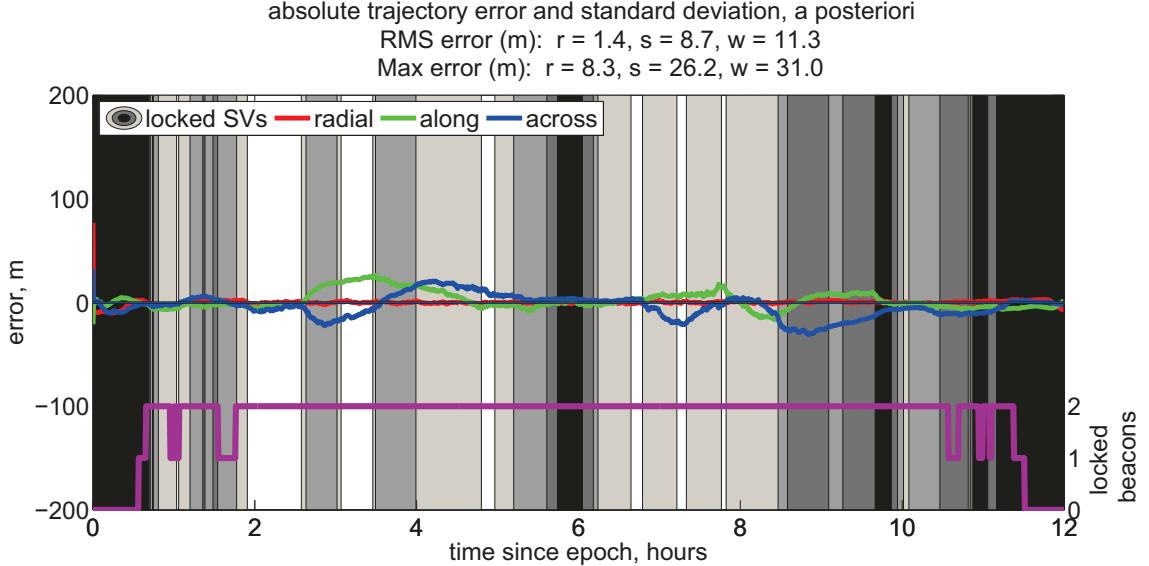


Figure 5.15: Navigation using passive ranging and GPS, receiver clock perfect. Beacon visibility indicated by the purple line at the bottom of the plot. The colour-code for the locked SVs is defined as before (cf. Figure 5.14).

our baseline simulation from Section 5.3.2, we note that despite the inclusion of new and accurate measurements, the accuracy of our navigation solution has not significantly improved. In particular, the radial coordinate is consistently under-estimated throughout the apogee arc. A comparison with the receiver clock error (not shown) confirms the same relationship we observed in Section 5.4.4. Namely, the filter compensates an under-estimation of the radial coordinate with an over-estimation of receiver clock bias. This error is magnified for the ground beacons because their line-of-sight to our receiver is very nearly aligned with the receiver's radial coordinate, reducing the filter's ability to distinguish between the two states. Thus, while the residuals remain small and all states remain within their respective variances, the solution is not optimal.

In an attempt to improve the solution, we increased the value of R_{rge} used in the filter while maintaining the measurement noise of $\sigma_{rge} = 5\text{ m}$. This forces the filter to lessen its dependence on the ranging measurements. Through a tuning process, a value of $R_{rge} = (50\text{ m})^2$ was found to be satisfactory. The performance of the filter with this value can be seen in Figure 5.17. Comparing its 3D RMS error of 29.1 m with previous simulations, we see a significant improvement over both the baseline simulation (58.9 m) and the pre-tuning beacon simulation (54.3 m). Thus, we have verified the inclusion of passive ranging as a means to improve our navigation solution.

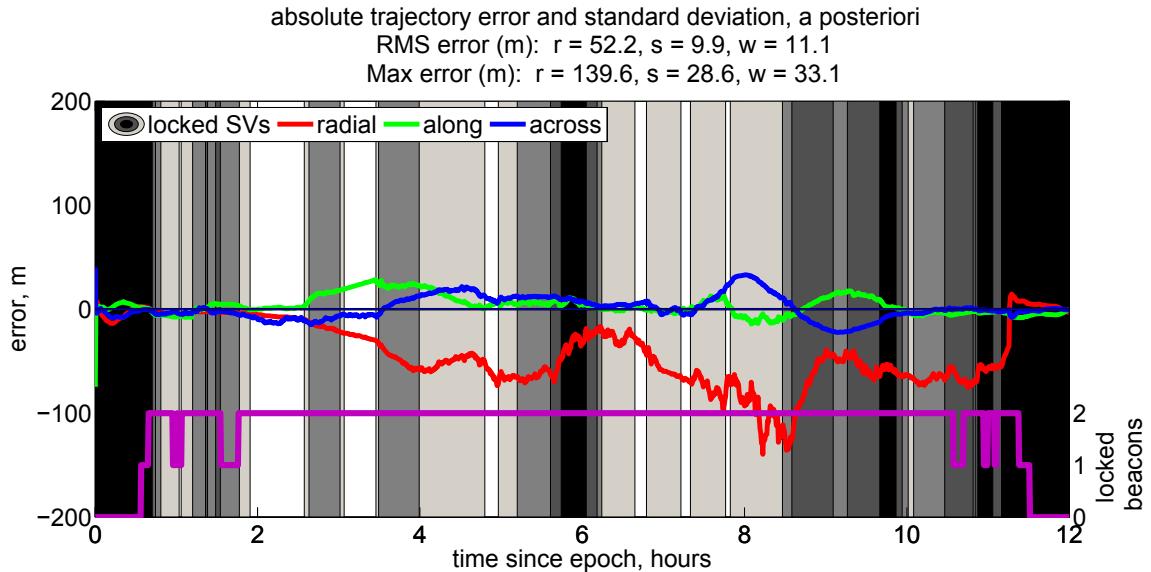


Figure 5.16: Navigation using passive ranging and GPS, receiver clock error included, $R_{rge} = (5 \text{ m})^2$.

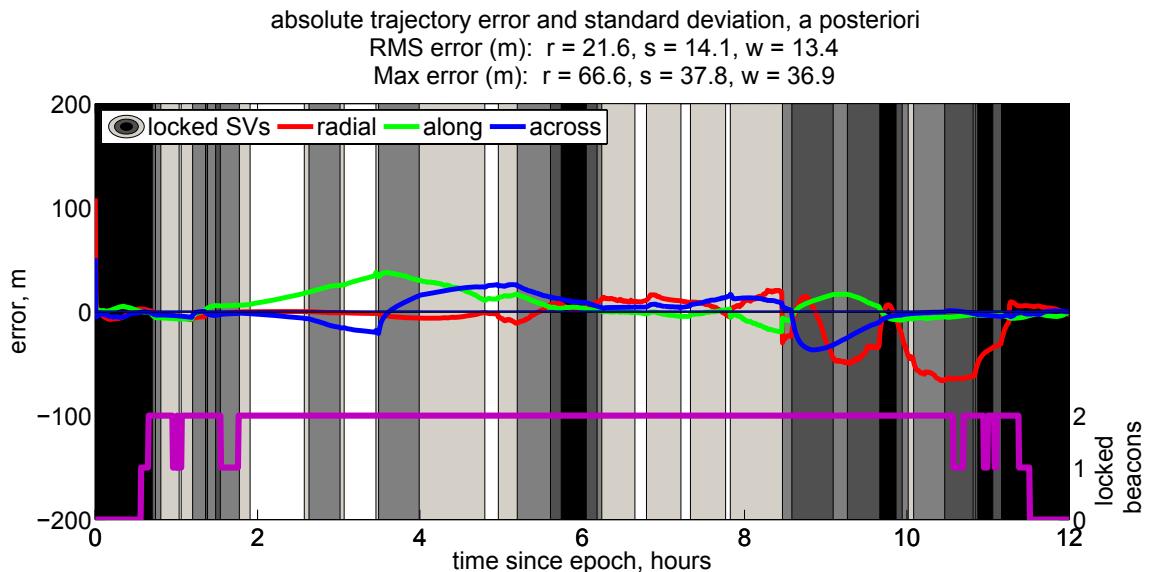


Figure 5.17: Navigation using passive ranging and GPS, receiver clock error included, $R_{rge} = (50 \text{ m})^2$.

5.5 Other Nonlinear Filters

As noted in Section 5.2, the Extended Kalman Filter (EKF) is the default filter for our state estimator. Thus, it has formed the backbone of our analysis so far. In Section 4.2, we introduced a quartet of alternatives to the EKF: the Unscented Kalman Filter, the Cubature Kalman Filter, the Continuous-Discrete Unscented Kalman Filter, and the Continuous-Discrete Cubature Kalman Filter. We now turn our attention to their implementation and evaluate their performance against the EKF.

For proper implementation of the Unscented Kalman Filter, a second round of filter tuning is required. In Section 4.2.2, we noted that the scaled unscented transform carries with it a trio of user-settable parameters: α , κ , and β . As outlined by Wan and van der Merwe [93]:

- α controls the spread of the sigma points around our state estimate $\hat{\mathbf{x}}$; it is usually set to a small positive value.
- κ is a secondary scaling parameter; it is typically set to 0 or $3 - n$, where n is the number of states in $\hat{\mathbf{x}}$.
- β allows us to emphasize or de-emphasize the weighting of the central covariance; it is typically set to 2 for a Gaussian distribution.

With these directives in mind, a tuning process led to the values seen in Table 5.2.

Table 5.2: Tuning parameters for the Unscented Kalman Filter

α	1×10^{-2}
κ	-5
β	2

The results of our tuned UKF simulation can be seen in Figure 5.18. In this figure, we can qualitatively see that the trajectory follows a similar pattern to that in our previous EKF-based simulations (see, e.g., Figure 5.5). The 3σ covariance bounds (not shown) behave likewise.

In a similar fashion, we applied each of the five filters to the same measurement dataset; their quantitative performance can be found in Table 5.3. Note that these

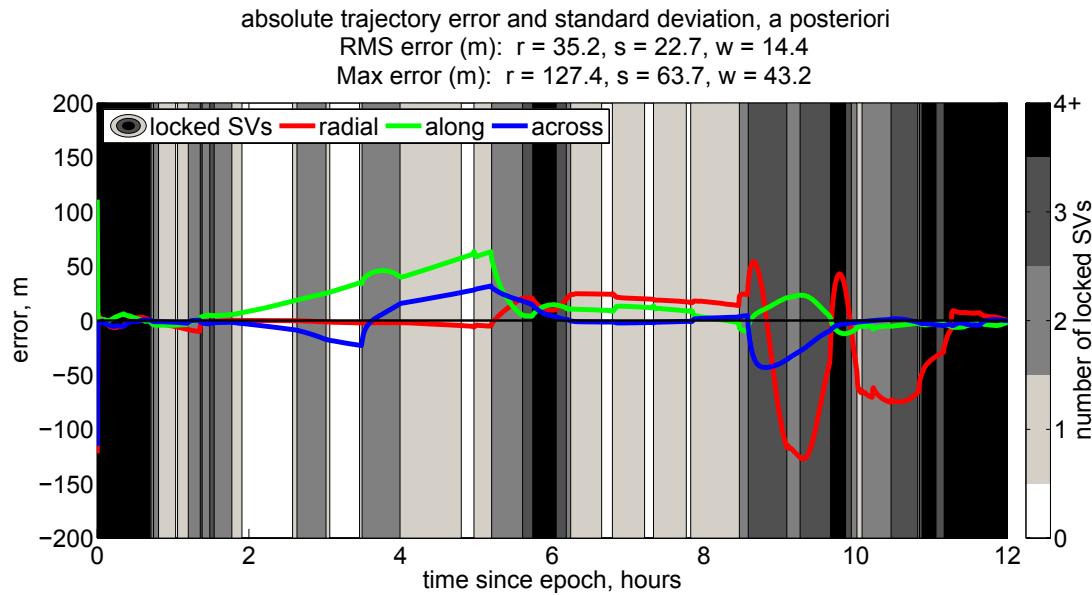


Figure 5.18: Unscented Kalman Filter, position error and locked GPS satellites

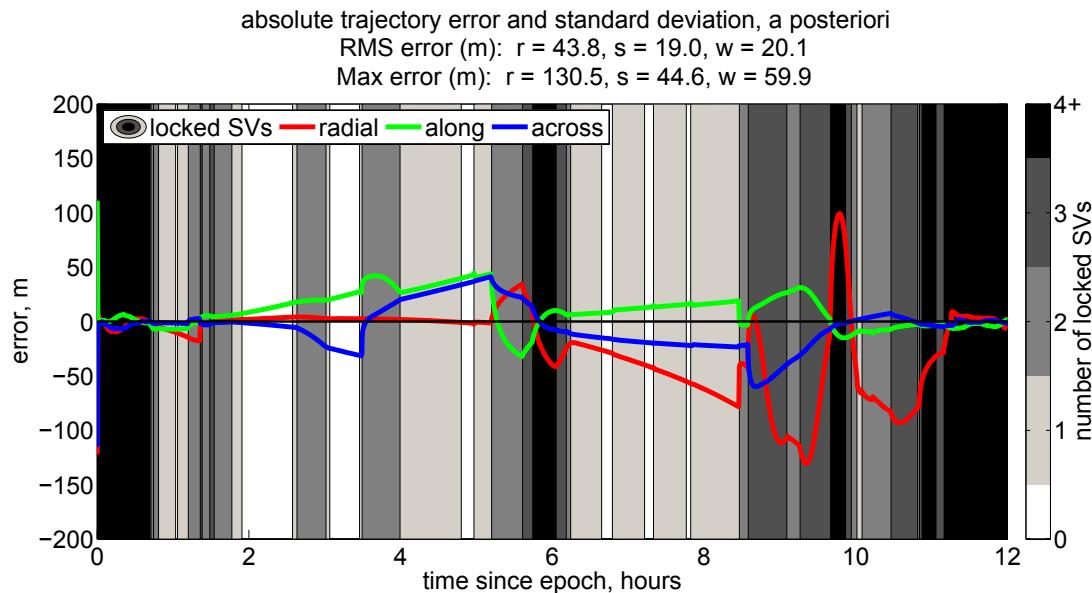


Figure 5.19: Continuous-Discrete Unscented Kalman Filter, position error and locked GPS satellites

Table 5.3: Summary of Results for Nonlinear Filters

Filter algorithm	Radial error	Along error	Across error	3D RMS error	Computational cost, relative
EKF	34.9	22.4	14.3	43.9	1
UKF	35.2	22.7	14.4	44.3	
CKF	35.2	22.7	14.4	44.3	
CDUKF	43.8	19.0	20.1	51.8	
CDCKF	34.6	22.4	14.2	43.6	

POSITION (m)	VELOCITY (mm·s ⁻¹)	EKF	UKF	CKF	CDUKF	CDCKF
		6.12	3.69	3.07	7.78	1
		6.08	3.70	3.08	7.76	10
		6.08	3.70	3.08	7.76	10
		7.51	3.95	3.82	9.31	15
		6.15	3.68	3.05	7.79	15

values cannot be compared directly to the values from the baseline case in Section 5.3.2 as they were conducted under different clock and ionospheric conditions (both are driven by random values).

Examining this table, we note that the accuracy of four of the five filters match within a metre. The outlier, the CDUKF, is plotted in Figure 5.19 for comparison. Though this trajectory is slightly different than the other four (which are indistinguishable from Figure 5.18), it retains a similar level of accuracy and is well within the operational requirements of our motivating PCW mission. The CDUKF result also serves to highlight that the same tuning for a discrete and continuous filter does not necessarily produce the same result.

The rightmost column of the table gives an indication of the computational time required for each filter when applied to our navigation problem. Exact computation times varied from run to run, but in general EKF analyses were executed in a tenth to fifteenth of the time of the other filters. This was due to two factors. First, the EKF algorithm requires only a single propagation of the nonlinear dynamics (cf. (4.8a)), while the UKF and CKF propagate these dynamics (including the computationally

intensive gravity model) for each sigma point (cf. (4.17c)). Second, as discussed in Section 5.2.2 a highly simplified Jacobian can be used in this application, further simplifying the EKF algorithm.

We recall from Section 4.3 that the UKF was not found to provide a significant improvement when implemented for GPS-based navigation in LEO. Further, we recall that the UKF and CKF were not shown to significantly improve on the EKF for ground-based position determination. Based on the results in Table 5.3, it appears that these conclusions hold for GPS-based navigation in HEO. Combining this conclusion with that of the rightmost column, we conclude that for the problem of GPS-based position determination in HEO, the Extended Kalman Filter is the computationally efficient and accurate solution.

5.6 Chapter Summary

In this chapter, we have summarized the verification of our GPS-based navigation solution for a satellite on highly elliptical orbit. First, we outlined our method of simulation and detailed our simulation configuration, including specific points about the propagator and measurement setup. Next, we defined and executed a baseline algorithm, which achieved a navigation accuracy within the requirements set by this study. Finally, we examined a variety of different filter configurations in an effort to optimize the navigation solution; these configurations are summarized in Table 5.4.

Having completed these analyses, we have demonstrated the ability of an EKF-based state estimator to navigate a spacecraft on highly elliptical orbit. We have also demonstrated that for this particular problem, estimators based on the UKF and CKF do not provide any additional accuracy despite their ability to better preserve system nonlinearities.

Table 5.4: Filter Design: Decisions and Observations

Design Feature	Options	Decision
receiver antenna	nadir-pointing zenith-pointing both	both
filter propagator	simple baseline complex	baseline
filter	EKF UKF CKF CDUKF CDCKF	EKF

Design Feature	Observation*
GPS sidelobes	improve the navigation solution
receiver clock error	most significant navigation error
passive ranging	improve the navigation solution

*further conclusions cannot be drawn without more detailed study.

Chapter 6

Conclusions

In this study, we have examined methods of nonlinear filtering for the autonomous navigation of spacecraft in highly elliptical orbits. This study was motivated by the upcoming Polar Communication and Weather mission, a Canadian Space Agency satellite slated for launch into a Molniya orbit. The satellite navigation requirement is specified at 150 m (3σ) in each component of an along/across/radial reference frame. Bearing this requirement in mind, we endeavoured to determine the accuracy to which such navigation can be achieved.

Defining our spacecraft state as its position and velocity, we first outlined the physical force models required to propagate this state forward in time. Next, we introduced measurements as a means of observing spacecraft state and presented detailed models of measurements that would be available to a spacecraft in HEO. We then summarized a number of nonlinear filtering algorithms designed to combine state predictions and measurements to produce a state estimate. Finally, we designed and executed a simulation to test these filter algorithms.

Our primary filtering algorithm, the Extended Kalman Filter, achieved a positional accuracy of 58.9 m (3D RMS) over a single Molniya orbit. Over this time, no single component of error exceeded 130 m; this satisfies the performance specification for our reference mission. Other nonlinear filters were not found to noticeably improve on this performance.

In the remainder of this chapter, we will present and discuss detailed conclusions, specify the contributions of this thesis, and outline possibilities for future work in this area of study.

6.1 Conclusions and Discussion

Orbit determination using GPS pseudoranges in HEO is “not nonlinear enough” to merit complex nonlinear filters. As noted above, our filtering simulation examined the performance of other nonlinear filters; these included the Unscented Kalman Filter, the Cubature Kalman Filter, and their continuous-discrete counterparts. These filters have been demonstrated to outperform the EKF for highly nonlinear systems and for long measurement outages. Nominally, the problem of orbit determination in HEO contains both such situations. However, these complex filters were not found to improve the navigation solution. When the dynamics and measurements are at their most nonlinear (at perigee due to strong geoidal gravity and rapidly changing geometry) GPS coverage is excellent, allowing the measurement update to quickly correct any propagation errors. When the measurements are infrequent (at apogee), the dynamics and measurements have limited nonlinearities; specifically, the Earth gravity approaches a point source and the pseudorange measurement approaches a linear combination of the radial coordinate and the receiver clock bias. Thus, for the measurements and sample periods considered in this thesis, a simple EKF with simple force models is sufficient.

When only one or two GPS pseudoranges are being processed, the navigation error is constrained but is noticeably worse than the error when three or more pseudoranges are available. In this study, great care was taken to implement realistic measurement errors for the GPS pseudorange signals. These errors included biased components (i.e., GPS SV ephemeris errors), slowly varying components (i.e., ionospheric errors), and random components (i.e., receiver clock noise). The tuned filter is able to reduce the residuals when one or two such measurements are processed, constraining the state measurement error but with a noticeably larger bound than the case when three or more pseudoranges are available. Thus, we have gained insight into the behaviour of a pseudorange-driven state estimator when GPS coverage is poor.

Receiver clock bias is a major source of error and needs to be treated with great care. At present, receiver clocks for spaceborne GPS satellites are prone to inaccuracy and require good GPS signal coverage to correct any biases that develop over time. Our HEO spacecraft is not able to fulfill this requirement and so must

estimate the receiver clock bias using sporadic GPS signals which are themselves corrupted by this bias. In this study, a dual-mode receiver clock model was shown to constrain receiver clock bias with just a single measurement. However, it has difficulty differentiating radial position error and receiver clock bias error as these states have similar impacts on the measurements estimated by the filter. We have therefore highlighted the need to examine the detailed behaviour of receiver clock bias in spaceborne receivers and the need for more refined techniques of bias estimation.

Passive ranging is an appealing option for autonomous satellite navigation in HEO. In this thesis, we introduced the concept of passive ranging as a means of providing additional autonomous measurements to our spacecraft. Briefly, this concept would require a pair of beacons set up in the Northern Hemisphere broadcasting a signal that is operationally similar to that broadcast by the GPS constellation. An exploratory study of this concept showed a 50% reduction in the position error of our navigation solution over the entire Molniya orbit.

More GPS signals result in a better solution. While this may seem to be a tautology, this study demonstrated that additional GPS signals unequivocally improved the navigation solution, particularly over the apogee arc where signal coverage is sparse. These additional signals can be facilitated by the inclusion of additional GPS antennae (i.e., two is better than one) and by using receivers and antennae capable of acquiring GPS sidelobe signals. Future missions in HEO would do well to prioritize the acquisition of GPS signals, including the optimization of antenna position for a mission-specific attitude profile.

6.2 Contributions

Based on the above, we can state the following specific contributions:

- well-founded **GPS pseudorange error modelling** to a high fidelity, including effects (GPS SV clock errors, GPS SV ephemeris errors) not previously included in HEO navigation studies;
- a thorough **filter and propagator comparison** which demonstrates that a simple filter and propagator are sufficient for GPS-based navigation in HEO;

- the demonstrated positive impact of **sidelobes and multiple GPS antennae** as means to acquire additional GPS signals;
- the significant impact of **receiver clock bias error** on the navigation solution and the filter's difficulty in differentiating between this error and a radial position error; and
- the introduction of **passive ranging** as a beneficial alternative measurement for HEO navigation.

6.3 Future Work

The focus of this study, autonomous navigation, was a driving factor behind the measurements and algorithms chosen. However, the constraint of **flight code** and on-orbit processing power was not considered. Future studies could further examine the outlined algorithms and determine their suitability for flight use in HEO.

Another option for future work is the use of an **augmented state vector** which includes empirical accelerations, force model parameters (i.e, drag coefficient and cross-section), and channel-specific GPS biases. The former would improve the propagator performance while the latter would act to counter the convergence issue detailed above. As such a state vector would increase computing requirements, additional study would be required to ensure the algorithm is appropriate for flight.

While this study focused primarily on orbit determination based on GPS pseudorange, there exist **other measurements** which have potential to improve the solution. Chief among these are the GPS carrier phase measurement (suitable for the elimination of the ionosphere pseudorange error) and accelerometer measurements (suitable for state estimation during thrusting events). In addition, certain LEO navigation technologies (i.e., the DORIS ground beacon system and the GLONASS satellite navigation system) could be verified for HEO through simulation or demonstration.

As noted above, the physical forces that define a Molniya orbit simplify as the orbital altitude increases; the Earth begins to resemble a point gravitational source and atmospheric drag becomes negligible. An **altitude-dependent process noise model** could be developed to ensure that the appropriate level of process noise is added to our filtering algorithm at all points in the orbit.

Finally, this study explicitly focused on the Molniya orbit. Further study could

extend these techniques to **other HEOs**, including the Tundra orbit and the Triple Apogee orbit.

References

- [1] *National Geographic Atlas of the World*. Washington, DC: National Geographic Society, 8th edition, 2004.
- [2] Statistics Canada. Canada's total population estimates, 2013. *The Daily*, component of Statistics Canada catalogue no. 11-001-X, 26 Sept 2013.
- [3] Edward Struzik. Alert. In *The Canadian Encyclopedia*. Historica Foundation, 2013.
- [4] Harry Kowalik. Telesat: The first domestic satellite system. In D.K. Sachdev, editor, *Success Stories in Satellite Systems*, page 93. Reston, VA: American Institute of Aeronautics and Astronautics, 2009.
- [5] Environment Canada. Satellite images and animation, accessed 20 Oct 2013. URL http://weather.gc.ca/satellite/index_e.html.
- [6] Canada. Governor General. *Seizing Canada's moment: prosperity and opportunity in an uncertain world: speech from the Throne, October 16, 2013*. Electronic serial in PDF format: <http://www.speech.gc.ca>.
- [7] Canada. Canadian Space Agency. Polar Communication and Weather mission (PCW), accessed 20 Oct 2013. URL <http://www.asc-csa.gc.ca/eng/satellites/pcw/>.
- [8] Alexander P. Trishchenko and Louis Garand. Canadian Polar Communication and Weather (PCW) satellite system: new capabilities for mapping Arctic snow and ice dynamics from highly elliptical orbit. 6th EARSeL Workshop: Cryosphere, Hydrology & Climate Interactions, University of Bern, Switzerland, 7 Feb 2011.
- [9] BBC News. Warming 'opens Northwest Passage', 14 Sept 2007. URL <http://news.bbc.co.uk/2/hi/6995999.stm>.
- [10] Anton H.J. de Ruiter, Christopher J. Damaren, and James R. Forbes. *Spacecraft Dynamics and Control*. West Sussex: John Wiley & Sons, Ltd., 2013.
- [11] Michael J. Bruno and Henry J. Pernicka. Tundra constellation design and stationkeeping. *Journal of Spacecraft and Rockets*, 42(5):902–912, 2005.

- [12] Alexander P. Trishchenko, Louis Garand, and Larisa D. Trichtchenko. Three-apogee 16-h highly elliptical orbit as optimal choice for continuous meteorological imaging of polar regions. *Journal of Atmospheric and Oceanic Technology*, 28:1407–1422, 2011.
- [13] David G. Hoag. Apollo navigation, guidance, and control systems: a progress report. Technical report, Cambridge, MA: MIT Instrumentation Library, April 1969.
- [14] Martin Unwin and Martin Sweeting. A practical demonstration of low cost autonomous orbit determination using GPS. In *Proceedings of the 8th International Technical Meeting of the Satellite Division of the Institutite of Navigation*, pages 579–587, Sept 1995.
- [15] Erin Kahr, Oliver Montenbruck, Kyle O’Keefe, Susan Skone, Jakub Urbanek, Laura Bradbury, and Pat Fenton. GPS tracking on a nanosatellite – the CanX-2 flight experience. In *Proceedings of the 8th International ESA Conference on GNC Systems*, Karlovy Vary, June 2011.
- [16] Michael L. Thrall. *Orbit Determinationa of Highly Eccentric Orbits using a RAVEN Telescope*. Master’s thesis, Naval Postgraduate School, 2005.
- [17] Dipak K. Srinivasan, David Artis, Ben Baker, Robert Stilwell, and Robert Wallis. RF communications subsystem for the Radiation Belt Storm Probes mission. *Acta Astronautica*, 65(11–12):1639–1649, 2009. doi: <http://dx.doi.org/10.1016/j.actaastro.2009.04.018>.
- [18] Jeff A. Estefan. Precise orbit determination of high-earth elliptical orbiters using differenced doppler and ranging measurements. In *IEEE Position Location and Navigation Symposium 1992*, pages 105–112, 1992.
- [19] George Davis, Michael Moreau, Russell Carpenter, and Frank Bauer. GPS-based navigation and orbit determination for the AMSAT AO-40 satellite. In *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Monterey, CA, Aug 2002.
- [20] James D. Kronman. Experience using GPS for orbit determination of a geosynchronous satellite. In *Proceedings of ION GPS 2000*, pages 1622–1626, Salt Lake City, UT, Sept 2000.

- [21] ESA Space Engineering. Far-out space navigation from sideways satnav signals, accessed 05 April 2014. URL http://www.esa.int/Our_Activities/Space_Engineering/Far-out_space_navigation_from_sideways_satnav_signals.
- [22] Rudolf E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME, Journal of Basic Engineering*, 82(D):35–45, 1960.
- [23] Rémi Kaupp. “Loch à plateau”, retrieved from http://commons.wikimedia.org/wiki/File:Loch_%C3%A0_plateau.jpg. Used under Creative Commons Attribution-ShareAlike 3.0 Unported license (<http://creativecommons.org/licenses/by-sa/3.0/deed.en>), 2007.
- [24] Oliver Montenbruck. An epoch state filter for use with analytical orbit models of low Earth satellites. *Aerospace Science and Technology*, 4(4):277–287, June 2000.
- [25] Isaac Newton. *Philosophiae Naturalis Principia Mathematica*. London, 1687. Translated by Ian Bruce, updated 26 Feb 2012, accessed 19 Nov 2013. URL <http://17centurymaths.com/contents/newtoncontents.html>.
- [26] David A. Vallado. *Fundamentals of Astrodynamics and Applications*. Hawthorne, CA: Microcosm Press, 3rd edition, 2007.
- [27] Iginio Marson. A short walk along the gravimeters path. *International Journal of Geophysics* 2012, 2012.
- [28] Richard H. Rapp. Past and future developments in geopotential modeling. In *Geodesy on the Move*, volume 119 of *International Association of Geodesy Symposia*, pages 58–78. Berlin: Springer-Verlag, 1998.
- [29] Helmut Groemer. *Geometric Applications of Fourier Series and Spherical Harmonics*. Cambridge: Cambridge University Press, 1996.
- [30] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions*. Washington, DC: National Bureau of Standards, 1972. Tenth Printing.
- [31] F.G. Lemoine, S.C. Kenyon, J.K. Factor, R.G. Trimmer, N.K. Pavlis, D.S. Chinn, C.M. Cox, S.M. Klosko, S.B. Luthcke, M.H. Torrence, Y.M. Wang, R.G.

- Williamson, E.C. Pavlis, R.H. Rapp, and T.R. Olson. The development of the joint NASA GSFC and NIMA geopotential model EGM96. Technical Report NASA/TP-1998-206861, NASA Goddard Space Flight Center, July 1998.
- [32] Matthew M. Berry and Vincent T. Coppola. Correct modeling of the indirect term for third-body perturbation. White paper, Analytical Graphics Inc., Aug 2007.
- [33] Steven Vogel. *Life In Moving Fluids: The Physical Biology of Flow*. Princeton, NJ: Princeton University Press, 1994.
- [34] James R. Wertz and Wiley J. Larson. *Space Mission Analysis and Design*. Torrance, CA: Microcosm Press, 3rd edition, 1999.
- [35] Oliver Montenbruck, Markus Markgraf, Joris Naudet, Stefano Santandrea, Kristof Gantois, and Pierrik Vuilleumier. Autonomous and precise navigation of the PROBA-2 spacecraft. In *Proceedings of the AIAA/AAS Astrodynamics Specialists Conference*, Honolulu, Aug 2008.
- [36] Dennis D. McCarthy and Gérard Petit. IERS conventions (2003). IERS Technical Note 32, International Earth Rotation and Reference Systems Service, 2004.
- [37] Heiner Klinkrad and Bent Fritzsche. Orbit and attitude perturbations due to aerodynamics and radiation pressure. In *Proceedings of the ESA Workshop on Space Weather*, ESTEC, Nov 1998.
- [38] Warren Soh. HEO STDP Navigation Trade-Off & Feasibility Study. Technical Report ER103975/B, Magellan Aerospace, Winnipeg, Oct 2012.
- [39] Derek J. de Solla Price. Contra-Copernicus: a critical re-estimation of the mathematical planetary theory of Ptolemy, Copernicus, and Kepler. In Marshall Clagett, editor, *Critical Problems in the History of Science*, pages 197–218. Madison, WI: University of Wisconsin Press, 1959.
- [40] Adrian Jäggi, Heike Bock, Lars Prange, Ulrich Meyer, and Gerhard Beutler. GPS-only gravity field recovery with GOCE, CHAMP, and GRACE. *Advances in Space Research*, 47(6):1020–1028, Mar 2011.

- [41] Keric A. Hill. *Autonomous Navigation in Libration Point Orbits*. PhD thesis, University of Colorado, Boulder, 2007.
- [42] Moriba K. Jah, Michael E. Lisano, George H. Born, and Penina Axelrad. Mars aerobraking spacecraft state estimation by processing inertial measurement unit data. *Journal of Guidance, Control, and Dynamics*, 31(6), 2008.
- [43] Dan Bradbury. What goes around comes around. *Engineering & Technology*, 8 (9):62–66, Oct 2013.
- [44] Ivan A. Getting. The Global Positioning System. *IEEE Spectrum*, 30(12):36–47, 1993.
- [45] United States Department of the Navy. The Tactical Use of Radar in Aircraft. Radar Bulletin No. 2 A, Office of the Chief of Naval Operations. Washington, DC: United States Government Printing Office, 1946.
- [46] Peter H. Dana. GPS time dissemination for real-time applications. *Real-Time Systems*, 12(1):9–40, Jan 1997.
- [47] U.S. Department of Homeland Security. GPS NANUs, Almanacs, & Ops Advisories, accessed 09 Dec 2013. URL <http://www.navcen.uscg.gov/?pageName=gpsAlmanacs>.
- [48] Global Positioning Systems Directorate. Navstar GPS Space Segment / Navigation User Interfaces. Technical Report IS-GPS-200G, Sept 2012.
- [49] Robert J. Murgatroyd. Troposphere. In *AccessScience*. McGraw Hill Education, 2012. URL <http://www.accessscience.com/content/troposphere/712200>.
- [50] Virgílio B. Mendes and Richard B. Langley. Tropospheric zenith delay prediction accuracy for airborne GPS high-precision positioning. In *Proceedings of the Institute of Navigation 54th Annual Meeting*, pages 337–347, 1998.
- [51] Michael C. Kelley and Francis S. Johnson. Ionosphere. In *AccessScience*. McGraw Hill Education, 2012. URL <http://www.accessscience.com/content/ionosphere/352700>.

- [52] John A. Klobuchar. Ionospheric time-delay algorithm for single-frequency GPS users. *IEEE Transactions on Aerospace and Electronic Systems*, 3:325–331, 1987.
- [53] International GNSS Service. IGS Products, accessed 02 April 2014. URL <http://www.igs.org/components/prods.html>.
- [54] Kenneth L. Senior, Jim R. Ray, and Ronald L. Beard. Characterization of periodic variations in the GPS satellite clocks. *GPS Solutions*, 12:211–255, 2008.
- [55] Micheal Moreau, Bo Naasz, Jesse Leitner, J. Russell Carpenter, and David Gaylor. Hardware in-the-loop demonstration of real-time orbit determination in high Earth orbits. In *Proceedings of the ION National Technical Meeting*, San Diego, Jan 2005.
- [56] Nicolay Mikhailov and Mikhail Vasil'ev. Autonomous satellite orbit determination using spaceborne GNSS receivers. *Gyroscopy and Navigation*, 2(1):1–9, 2011.
- [57] Oliver Montenbruck. Performance assessment of the NovAtel OEM4-G2 receiver for LEO satellite tracking. Technical Report TN 03-05, Space Flight Technology, DLR/GSOC, July 2003.
- [58] Benjamin Dainty, John Raquet, and Richard Beckman. Improving geostationary satellite GPS position error using dynamic two-way time transfer measurements. In *Proceedings of the 39th Annual Precise Time and Time Interval Meeting*, pages 511–530, Nov 2007.
- [59] Lawrence R. Weill. GNSS Solutions: differences between signal acquisition and tracking. *InsideGNSS*, 6(1), 2011.
- [60] Jorge Potti, Miguel Bello, and Alessandro Pasotti. A GPS-based orbit positioning system for the European Archimedes mission. In *Proceedings of the Third International Conference on Spacecraft GNC Systems*, pages 67–75, Nov 1996.
- [61] Warren Soh, Mark Swartz, and Tarun Patel. HEO AODCS Simulator Development and Validation Report. Technical Report ER104011, Magellan Aerospace, Winnipeg, Mar 2013.

- [62] Shankararaman Ramakrishnan, Tyler Reid, and Per Enge. Leveraging the common GPS and Galileo L1 composite signal to enable autonomous navigation at GEO and beyond. In *26th International Technical Meeting of the Satellite Division of the Institute of Navigation*, Sept 2013.
- [63] José Lorga, Pedro Silva, Fabio Dovis, Andrea Di Cintio, Steeve Kowaltschek, David Jiminez, and Roger Jansson. Autonomous orbit determination for future GEO and HEO missions. In *5th ESA Workshop on Satellite Navigation Technologies*, pages 1–14, Dec 2010.
- [64] James Kronman. Experience using GPS for orbit determination of a geosynchronous satellite. In *Proceedings of the Institute of Navigation GPS 2000 Conference*, pages 1622–1626, Sep 2000.
- [65] Michael C. Moreau, Edward P. Davis, J. Russell Carpenter, David Kelbel, George W. Davis, and Penina Axelrad. Results from the GPS flight experiment on the high Earth orbit AMSAT OSCAR-40 spacecraft. In *Proceedings of the Institute of Navigation GPS 2002 Conference*, pages 122–133, Sep 2002.
- [66] Daryl Mitchell. “SED Systems”, retrieved from http://commons.wikimedia.org/wiki/File:SED_Systems.jpg. Used under Creative Commons Attribution-ShareAlike 3.0 Unported license (<http://creativecommons.org/licenses/by-sa/3.0/deed.en>), 2006.
- [67] Yoaz Bar-Sever, Larry Young, Frank Stocklin, and John Rush. NASA’s global differential GPS system and the TDRSS augmentation service for satellites. In *Proceedings of the 2nd ESA Workshop on Satellite Navigation User Equipment Technologies*, ESTEC, Dec 2004.
- [68] Angela Reichert, Thomas Meehan, and Timothy Munson. Toward decimeter-level real-time orbit determination: a demonstration using the SAC-C and CHAMP spacecraft. In *Proceedings of the 15th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2002)*, Portland, OR, Sep 2002.
- [69] Christian Jayles, Jean-Pierre Chauveau, and Fabien Rozo. DORIS/Jason-2: Better than 10 cm on-board orbits available for near-real-time altimetry. *Advances in Space Research*, 46(4):1497–1512, 2010.

- [70] Sergey Klyushnikov, Sergey Filatchenkov, Nicolai Mikhailov, Sergey Pospelov, and Mikhail Vasilyev. “Soyuz” - “Mir” orbital flight GPS/GLONASS experiment: first results. In *Proceedings of the RTO SCI International Conference on Integrated Navigation Systems*, St. Petersburg, May 1999.
- [71] Oliver Montenbruck, Paul Swatschina, Markus Markgraf, Stefano Santandrea, Joris Naudet, and Etienne Tilmans. Precision spacecraft navigation using a low-cost GPS receiver. *GPS Solutions*, 16(4):519–529, 2012.
- [72] Renato Zanetti and Chris D’Souza. Dual accelerometer usage strategy for on-board space navigation. *Journal of Guidance, Control, and Dynamics*, 35(6):1899–1902, 2012.
- [73] Warren Soh and Ron Buckingham. HEO STDP Navigation Tundra Assessment. Technical Report ER104398, Magellan Aerospace, Winnipeg, Mar 2014.
- [74] Mohinder S. Grewal and Angus P. Andrews. Applications of Kalman Filtering in Aerospace 1960 to the Present. *IEEE Control Systems Magazine*, pages 69–78, Jun 2010.
- [75] John L. Crassidis and John L. Junkins. *Optimal Estimation of Dynamic Systems*. Boca Raton, FL: CRC Press LLC, 2nd edition, 2011.
- [76] Dan Simon. *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Hoboken, NJ: John Wiley & Sons, Inc., 1st edition, 2006.
- [77] NASA. “Apollo 10 Lunar Module Rendezvous”, retrieved from http://commons.wikimedia.org/wiki/File:Apollo_10_Lunar_Module_Rendezvous.jpg. Public domain image courtesy of NASA, 22 May 1969.
- [78] Jeffrey K. Uhlmann. First hand: the unscented transform, on *IEEE Global History Network*, 2012. URL http://www.ieeeghn.org/wiki/index.php/First-Hand:The_Unscented_Transform.
- [79] Simon J. Julier and Jeffrey K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, Mar 2004.
- [80] Soumyo Dutta and Robert D. Braun. Statistical entry, descent, and landing performance reconstruction of the Mars Science Laboratory. In *Proceedings of*

the AIAA Atmospheric Flight Mechanics Conference, National Harbor, MD, Jan 2014.

- [81] Eric A. Wan and Rudolph van der Merwe. The Unscented Kalman Filter for nonlinear estimation. In *Proceedings of the IEEE Symposium on Adaptive Systems for Signal Processing, Communications, and Control*, pages 153–158, Lake Louise, AB, Oct 2000.
- [82] Ienkaran Arasaratnam and Simon Haykin. Cubature Kalman filters. *IEEE Transactions on Automatic Control*, 54(6):1254–1269, 2009.
- [83] Simo Särkkä. On Unscented Kalman Filtering for state estimation of continuous-time nonlinear systems. *IEEE Transactions on Automatic Control*, 52(9):1631–1641, 2007.
- [84] Howard Heuberger and Leonard Church. Landsat-4 Global Positioning System navigation results. In *Proceedings of the AAS/AIAA Astrodynamics Conference, Part I*, pages 589–602, Aug 1983.
- [85] Eberhard Gill, Oliver Montenbruck, and Sergio Montenegro. Flight results from the BIRD onboard navigation system. In *Proceedings of the 5th International ESA Conference on GNC Systems*, Frascati, Oct 2002.
- [86] Jean-Sébastien Ardaens, Simone D’Amico, and Oliver Montenbruck. Final commissioning of the PRISMA GPS navigation system. In *Proceedings of the 22nd International Symposium on Spaceflight Dynamics*, São José dos Campos, Mar 2011.
- [87] Taohu Li, Jiansheng Liu, Zhigang Huang, and Honglei Qin. Observability of HEO satellite autonomous navigation system using GPS. In *Proceedings of the 2010 International Conference on Multimedia Technology*, Ningbo, Oct 2010.
- [88] Eun-Jung Choi, Jae-Cheol Yoon, Byoung-Sun Lee, Sang-Young Park, and Kyu-Hong Choi. Onboard orbit determination using GPS observations based on the unscented Kalman filter. *Advances in Space Research*, 46(11):1440–1450, 2010.
- [89] Henri Pesonen and Robert Piché. Cubature based Kalman filters for positioning. In *Proceedings of the 7th Workshop on Positioning, Navigation, and Communication*, Dresden, Mar 2010.

- [90] Peter Zentgraf, Stan Berge, Camille Chasset, Hannes Filippi, Eveline Gottzein, Ignacio Gutiérrez-Cañas, Mark Hartrampf, Peter Krauss, Christopher Kuehl, Bernhard Lübbe-Ossenbeck, Michael Mittnacht, Oliver Montenbruck, Carsten Müller, Pablo Rueda Boldo, and Attilio Truffi. Preparing the GPS-experiment for the Small GEO mission. In *Proceedings of the 33rd AAS Guidance and Control Conference*, Breckenridge, CO, Feb 2010.
- [91] Jean-Paul Berthias, Patrick Broca, Jean Fourcade, Christian Jayles, Didier Pradines, and Denis Laurichesse. General characteristics of real-time on-board orbit determination. In *Proceedings of the 12th International Symposium on Space Flight Dynamics*, pages 267–274, ESOC, Jun 1997.
- [92] United States Coast Guard. Navstar GPS user equipment introduction. Technical report, U.S. Coast Guard Navigation Center, Alexandria, VA, Sep 1996.
- [93] Eric A. Wan and Rudolph van der Merwe. Chapter 7: The Unscented Kalman Filter. In Simon Haykin, editor, *Kalman Filtering and Neural Networks*, pages 221–280. New York: John Wiley & Sons, Inc., 2001.
- [94] William M. Folkner, James G. Williams, and Dale H. Boggs. The planetary and lunar ephemeris DE 421. IPN Progress Report 42-178, Jet Propulsion Laboratory, Aug 2009.
- [95] Richard Langley. Dilution of precision. *GPS World*, pages 52–59, May 1999.
- [96] Oliver Montenbruck and Eberhard Gill. *Satellite Orbits: Models, Methods, and Applications*. Berlin: Springer-Verlag, 1st edition, 2000.
- [97] John C. Butcher. A history of Runge-Kutta methods. *Applied Numerical Mathematics*, 20:247–260, 1996.
- [98] MATLAB Documentation Centre. Solve nonstiff differential equations, medium order method - MATLAB ode45, accessed 08 April 2014. URL <http://www.mathworks.com/help/matlab/ref/ode45.html>.
- [99] Manuel Ruiz Delgado. Radiation pressure: Modelling the space environment. Presentation, Universidad Politécnica de Madrid, Apr 2008. URL http://ocw.upm.es/ingenieria-aeroespacial/modeling-the-space-environment/contenidos/material-de-clase/mse07_radiationpressure.pdf.

- [100] Robert B. Lee, III, M. Alan Gibson, Robert S. Wilson, and Susan Thomas. Long-term total solar irradiance variability during sunspot cycle 22. *Journal of Geophysical Research*, 100(A2):1667–1675, 1995.

Appendix A

Supplementary Mathematics

A.1 The Classical Orbital Elements

In Table 2.3, we defined the orbital initial conditions for this study using a set of parameters known as orbital elements. These elements were developed to describe the motion of a particle subject to the two-body acceleration. They are therefore an appropriate means to characterize a spacecraft in elliptical orbit about the Earth.

Each one of these elements has a specific physical interpretation. In brief:

- Two of these elements define the shape of the orbit. The **semi-major axis** a is half the distance across the long axis of the elliptical orbit. The **eccentricity** e is a unitless parameter that describes the ellipticity of the orbit, ranging from nearly zero (nearly circular) to nearly one (highly elliptical).
- Two of the elements define the plane of the orbit with respect to a fundamental plane. For Earth-orbiting satellites, this fundamental plane is the equatorial plane (i.e., that plane which is perpendicular to the Earth's spin axis). The **inclination** i gives the angle between the orbital plane and the equatorial plane. The **right ascension of the ascending node** Ω is an angle which defines the location of the ascending node (i.e., the point where the sub-satellite point crosses from the Southern to Northern Hemisphere) with respect to a standard reference direction, the First Point of Aries¹.
- A fifth parameter, the **argument of perigee** ω , defines the orientation of the

¹The First Point of Aries is the intersection of the Earth equatorial plane and Earth orbital plane around the Sun.

orbit in its plane by specifying the angular location of the perigee with respect to the ascending node.

- The final parameter is the **true anomaly** ν , an angle which defines the location of the satellite with reference to the orbit perigee at the epoch of interest t_0 .

For further details on these elements, please consult Section 2.4.1 of Vallado [26].

A.2 Reference Frames

Throughout this study, we have discussed and developed vector formulae describing the kinematics and kinetics of spacecraft navigation. For computational purposes, these vectors must be expressed in a reference frame. Spacecraft dynamics commonly relies on two broad categories of reference frames: Earth-Centred Inertial (ECI) frames and Earth-Centred Earth-Fixed (ECEF) frames.

An **ECI frame** is centred at the mass-centre of the Earth and is the standard frame for all calculations related to the equations of motion. Its axes are defined by two vectors: the x-axis points along the First Point of Aries while the z-axis is aligned with the Earth's rotation axis. Though this frame is defined by directions fixed in inertial space, it is *quasi-inertial* as its origin is moving with a non-constant velocity when referenced to the solar system barycentre (the nearest reasonable approximation to an inertially fixed point). However, for Earth-orbiting satellites, this slight distinction is inconsequential and Newtonian mechanics can be applied in an ECI frame as though it were truly inertial.

Like the ECI frame, an **ECEF frame** is also centred at the mass-centre of the Earth and is the standard frame for any phenomena explicitly tied to the shape of the Earth. Unlike the ECI frame, the ECEF axes are fixed to the figure of the Earth itself: the x-axis nominally corresponds to the crossing of the Equator and Prime Meridian (0°N , 0°E) while the z-axis nominally corresponds to the terrestrial North Pole (90°N). In this study, an ECEF frame is used for one specific reason: calculation of the geoidal acceleration (see Appendix B).

Since this study uses both ECI and ECEF frames, accurate conversion between them is essential. Briefly, this conversion must account for:

- **precession** and **nutation** of the Earth's rotation axis due to the gravity of the Sun and Moon pulling on our non-spherical Earth;

- the **rotation** of the Earth about its axis, repeating once every 24 h;
- the **polar motion** of the instantaneous Earth rotation axis when compared to the terrestrial North Pole.

As an illustrative example of the care that must be taken in this coordinate transformation, the final effect (polar motion) is an arcsecond-level adjustment which translates to a few metres of polar motion on the surface of the Earth. Yet, as discussed by Montenbruck et al. [35] this effect is essential to any Kalman-based navigation exercise to retain propagator accuracy – a conclusion that was re-affirmed in the development of the propagator for this study.

As discussed above, ECI and ECEF are general categories of reference frames. In this study, we used the J2000 frame of the 5th Fundamental Catalog (FK5) as our ECI frame and the 1993 version of the International Terrestrial Reference Frame (ITRF93) as our ECEF frame. Both are compatible with the JPL Development Ephemeris described in the next section.

For further discussion on reference frames, please consult Section 3.7 of Vallado [26].

A.2.1 The JPL Development Ephemeris

In support of ongoing spacecraft analyses, NASA’s Jet Propulsion Laboratory (JPL) maintains a highly accurate ephemeris for our Solar System. This Developmental Ephemeris (DE) is continuously improved as new measurements become available; recent editions incorporate data from dozens of missions and measurement campaigns over the past half-century [94]. Building off this proven track record, this study employed the DE 421 ephemeris for solar system ancillary data.

To access the ephemeris, JPL provides the SPICE toolkit, a C-based interface which is available in a MATLAB-compatible edition (MICE). By using the MICE toolkit, our MATLAB-based navigation simulation can access:

- **lunar and solar ephemerides**, necessary for calculation of the third-body acceleration (2.5) and the solar radiation pressure acceleration (2.8);
- the **ECI-ECEF frame rotation** necessary to apply the geoidal acceleration (see previous section);

- **time conversions** between UTC, GPS time, and the ephemeris time upon which the ephemeris is defined.

All of these parameters are calculated from static datafiles distributed as a part of the SPICE toolkit. These datafiles are accurate until the year 2050, well beyond the expected lifetime of our spacecraft. Thus, this is an appropriate resource for our navigation simulation.

A.3 Time Dilution of Precision

In Section 3.5.1, we noted that our receiver clock truth will only be steered when the Time Dilution of Precision (TDOP) is below a threshold of 10. In the following section, we will define TDOP and discuss the threshold chosen for our current application.

As discussed by Langley [95], the relative geometry between receivers and transmitters in a radionavigation system has a direct impact on the accuracy of the resulting point solution. The Geometric Dilution of Precision (GDOP) metric is a means by which to quantify this effect. Roughly, GDOP is a unitless value which relates input measurement uncertainty to point solution uncertainty:

$$GDOP = \frac{(\text{point solution uncertainty})}{(\text{measurement uncertainty})} \quad (\text{A.1})$$

Mathematically, the calculation of GDOP relies on an intermediate matrix \mathbf{A} describing the current receiver-SV geometry:

$$\mathbf{A} = \text{col}_i \begin{bmatrix} \hat{\mathbf{e}}_{\mathbf{r}_{sat \rightarrow SV(i)}}^T & -1 \end{bmatrix} \quad (\text{A.2})$$

where $\hat{\mathbf{e}}_{\mathbf{r}_{sat \rightarrow SV(i)}}$ is the unit vector from the receiver to each GPS SV for which the receiver currently has lock. When at least four SVs are locked, the GDOP matrix \mathbf{D} can be calculated as:

$$\mathbf{D} = (\mathbf{A}^T \mathbf{A})^{-1} \quad (\text{A.3})$$

where the values of interest are found along the main diagonal of \mathbf{D} . Together, these four values define the GDOP for a given receiver-SV geometry.

Because we are working with receiver clock bias, we are interested in the fourth

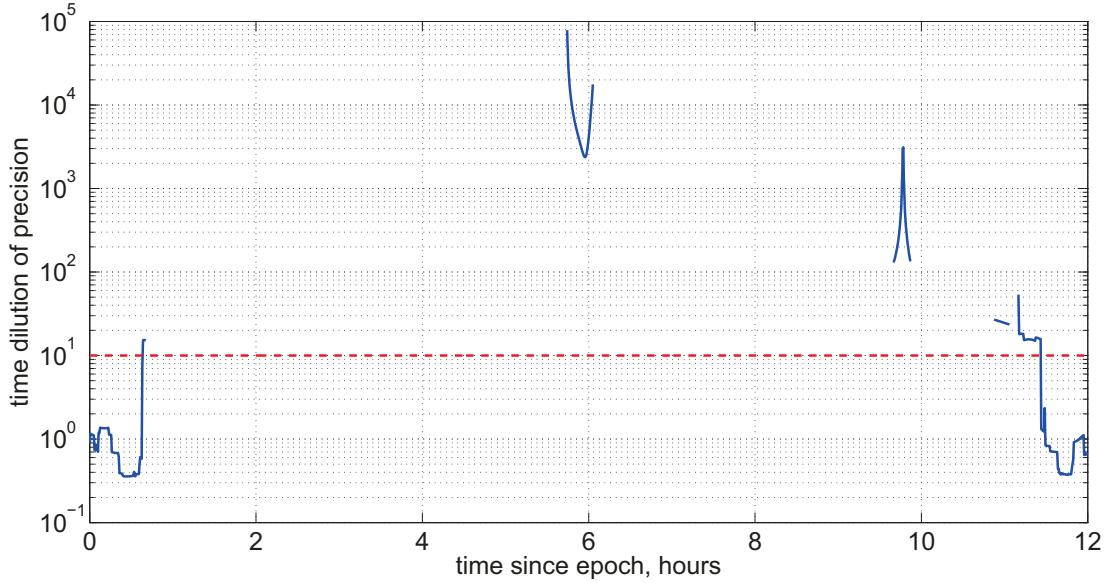


Figure A.1: Time Dilution of Precision over a single Molniya orbit. The red dashed line indicates the threshold $TDOP = 10$. Simulation begins at perigee.

component of this diagonal. TDOP is therefore calculated as:

$$TDOP = \sqrt{d_{44}} \quad (\text{A.4})$$

where d_{44} is an element of the matrix \mathbf{D} defined above.

A.3.1 TDOP for a Molniya orbit

The United States Coast Guard [92] notes that the dilution of precision will rise (i.e., the point solution accuracy degrades) when the line-of-sight unit vectors $\hat{\mathbf{e}}_{\mathbf{r}_{sat \rightarrow SV(i)}}$ are nearly aligned. As visualized in Figure 3.4, this near-alignment will occur when our spacecraft is at apogee and receiving multiple signals from SVs on the far side of the Earth. In an attempt to characterize the impact of the varying GPS geometry on our navigation solution, we calculated TDOP over one Molniya orbit. The results can be seen in Figure A.1.

TDOP at perigee: From the figure, we note that at perigee the TDOP hovers around unity. This is an expected result for a satellite with excellent GPS coverage

(cf. Figure 3.3). As the number of locked SVs falls, we see a corresponding rise in TDOP, continuing until less than four satellites are locked. At this point, a calculation of TDOP is meaningless (point solutions are not possible) and no further values are produced until apogee.

TDOP at apogee: As confirmed by Figure 3.3, there is a brief period at apogee where four SVs are locked by the receiver. All of these are located on the far side of the Earth. As their line-of-sight unit vectors are nearly aligned, the TDOP at this point is very high, indicating that metres of error in our pseudorange would correspond to tens of kilometres of point solution error. This corresponds to tens of microseconds of receiver clock error. As our receiver clock estimation algorithm has a demonstrated accuracy in the hundreds of nanoseconds (cf. Section 5.3.2), it would be detrimental to begin receiver clock steering using these point solutions.

A.3.2 Selecting a TDOP threshold

Drawing from the clock estimator performance in Figure 5.3 and the TDOP simulation in Figure A.1, a TDOP threshold of 10 was chosen for our clock-steering algorithm. This means that clock steering will only occur if the TDOP is less than 10, a threshold indicated in Figure A.1 by a red dashed line.

Given that our pseudoranges are typically accurate to within ten metres, a TDOP of 10 corresponds to clock accuracies in the hundreds of nanoseconds. As stated in the previous section, this is approximately the threshold of our receiver clock estimation algorithm. Thus, $TDOP = 10$ is an appropriate point at which to transition from a steered to drifting receiver clock.

Appendix B

Recursive Algorithm for Earth Gravity

As introduced in Section 2.2.1, there exist recursive algorithms for the evaluation of the acceleration predicted by a spherical harmonic model of the Earth's gravitational field. The algorithm implemented in this study is that given by Montenbruck and Gill [96] (p. 66–68); it is reproduced here for completeness.

First, we define:

$$V_{nm} = \left(\frac{R_\oplus}{r} \right)^{n+1} \cdot P_{nm}(\sin \phi) \cdot \cos(m\lambda) \quad (\text{B.1a})$$

$$W_{nm} = \left(\frac{R_\oplus}{r} \right)^{n+1} \cdot P_{nm}(\sin \phi) \cdot \sin(m\lambda) \quad (\text{B.1b})$$

allowing us to write the gravitational potential as

$$U = \frac{\mu_\oplus}{R_\oplus} \sum_{n=0}^{\infty} \sum_{m=0}^n (C_{nm} V_{nm} + S_{nm} W_{nm}) . \quad (\text{B.2})$$

Using the definition in (B.1), V_{nm} and W_{nm} satisfy the following recurrence relations:

$$V_{mm} = (2m - 1) \left\{ \frac{x R_\oplus}{r^2} V_{m-1,m-1} - \frac{y R_\oplus}{r^2} W_{m-1,m-1} \right\} \quad (\text{B.3a})$$

$$W_{mm} = (2m - 1) \left\{ \frac{x R_\oplus}{r^2} W_{m-1,m-1} + \frac{y R_\oplus}{r^2} V_{m-1,m-1} \right\} \quad (\text{B.3b})$$

$$V_{nm} = \left(\frac{2n - 1}{n - m} \right) \cdot \frac{z R_\oplus}{r^2} \cdot V_{n-1,m} - \left(\frac{n + m - 1}{n - m} \right) \cdot \frac{R_\oplus^2}{r^2} \cdot V_{n-2,m} \quad (\text{B.4a})$$

$$W_{nm} = \left(\frac{2n-1}{n-m} \right) \cdot \frac{zR_\oplus}{r^2} \cdot W_{n-1,m} - \left(\frac{n+m-1}{n-m} \right) \cdot \frac{R_\oplus^2}{r^2} \cdot W_{n-2,m} \quad (\text{B.4b})$$

with initial values

$$V_{00} = \frac{R_\oplus}{r} , \quad W_{00} = 0 . \quad (\text{B.5})$$

Thus, we can use (B.3), (B.4), and (B.5) to determine V_{nm} and W_{nm} to our desired degree and order. The accelerations are then directly calculated as

$$\ddot{x} = \sum_{n,m} \ddot{x}_{nm} , \quad \ddot{y} = \sum_{n,m} \ddot{y}_{nm} , \quad \ddot{z} = \sum_{n,m} \ddot{z}_{nm} \quad (\text{B.6})$$

where the partial accelerations are given by

$$\ddot{x}_{nm} \stackrel{(m=0)}{=} \frac{\mu_\oplus}{R_\oplus^2} \cdot \left\{ -C_{n0} V_{n+1,1} \right\} \quad (\text{B.7a})$$

$$\begin{aligned} &\stackrel{(m>0)}{=} \frac{\mu_\oplus}{R_\oplus^2} \cdot \frac{1}{2} \cdot \left\{ (-C_{nm} V_{n+1,m+1} - S_{nm} W_{n+1,m+1}) \right. \\ &\quad \left. + \frac{(n-m+2)!}{(n-m)!} \cdot (+C_{nm} V_{n+1,m-1} + S_{nm} W_{n+1,m-1}) \right\} \end{aligned} \quad (\text{B.7b})$$

$$\ddot{y}_{nm} \stackrel{(m=0)}{=} \frac{\mu_\oplus}{R_\oplus^2} \cdot \left\{ -C_{n0} W_{n+1,1} \right\} \quad (\text{B.7c})$$

$$\begin{aligned} &\stackrel{(m>0)}{=} \frac{\mu_\oplus}{R_\oplus^2} \cdot \frac{1}{2} \cdot \left\{ (-C_{nm} W_{n+1,m+1} + S_{nm} V_{n+1,m+1}) \right. \\ &\quad \left. + \frac{(n-m+2)!}{(n-m)!} \cdot (-C_{nm} W_{n+1,m-1} + S_{nm} V_{n+1,m-1}) \right\} \end{aligned} \quad (\text{B.7d})$$

$$\ddot{z}_{nm} = \frac{\mu_\oplus}{R_\oplus^2} \cdot \left\{ (n-m+1) \cdot (-C_{nm} V_{n+1,m} - S_{nm} W_{n+1,m}) \right\} . \quad (\text{B.7e})$$

Note that these accelerations are inertial but represented in an Earth-fixed reference frame. For details on reference frame transformations, see Appendix A.2.

Appendix C

Simulation Details

As discussed in Section 1.4, this study documents a simulation-based analysis of satellite navigation techniques. While the focus of this document has been the theory and equations underlying the simulation, it is prudent to briefly outline the process by which these equations were implemented.

The simulations outlined throughout this study involve three pieces of software:

- STK 9.0, an industry-standard environment for astrodynamical analyses;
- MATLAB R2010a, a mathematical software environment;
- Tapestry, purpose-built software for the LabPro-4000 GPS simulator.

Further details on Tapestry and MATLAB environment can be found in the remainder of this Appendix.

C.1 Using the GPS simulator

Note: The following is a minimal set of instructions that assumes a user is familiar with the operation of the LabPro-4000 GPS simulator. For further information, please consult the simulator documentation, including `GettingStarted.pdf`.

1. Open up Tapestry (GPS simulator software) and *Create A New Scenario* with 1 vehicle and 2 antennae.
2. Import a trajectory per `ImportTrajectoryFile.pdf` with desired truth orbit and attitude.

3. Configure GPS antenna per [UsingAntennaModels.pdf](#): gain pattern, boresight, and lever arm. Note that gain patterns are actually attenuation patterns where a value of 0.0 corresponds to the point of maximum gain.
4. Under *SV MENU / ANTENNA PATTERN OPTIONS*, toggle *ENABLE UE GAIN PATTERNS* to the checked state.
5. Configure environmental models per [RangeModels.pdf](#): troposphere (off), ionosphere (shell with default values), multipath (off), selective availability (off), and user equivalent range error (off).
6. Under *Utilities*, open the configuration textfile `tapcontrol.ini` and find the heading `[GPSCONTROL]`. Under this heading:
 - Set `SVReviewInternal=30`; this controls the frequency with which the GPS constellation is propagated (once every 30 seconds).
 - Set `SVMaskAngle=-90`; this sets the GPS receiver mask to 90° below the horizon, effectively permitting 4π SV visibility.

Both of these steps configure the simulator for spaceborne use.

7. Under *SV MENU* select *OPEN SATPOWER PRO*. Select GPS satellites 1–16 per [EditingSATpowerPro.pdf](#).
8. Save scenario by clicking *BUILD RF OUTPUT*.
9. Re-open scenario and click *Utilities / Dump Range Truth*. Select *All the Above* and save a file for RF#1 (first antenna) and RF#2 (second antenna). Select *ECEF Satellite State Vector* and save a single file. For further details, consult [DumpRange.pdf](#).
10. Repeat previous three steps for GPS satellites 17–32.

These steps will provide the information detailed in Section 3.6.1. The user must then calculate the SV antenna gain (a function of geometry) and check for Earth-blockage before a link budget can be calculated to check whether a given SV is visible to the receiver.

C.2 The MATLAB simulation environment

The MATLAB-based navigation simulation requires a number of ancillary files:

- measurement data products from the GPS Simulator (see Section 3.6.1);
- ancillary astrodynamical information from the JPL Developmental Ephemeris (see Appendix A.2), including solar/lunar ephemerides and coordinate frame transformations;
- the definition of the GPS SV antenna¹;
- datafiles defining the original truth orbit from STK, the EGM-96 coefficients, and the errors in the broadcast GPS navigation message as listed in Appendix E.

These files are saved as text files in a directory known to the simulation and are loaded into the simulation using MATLAB’s `textread` function.

Form of the numerical propagator: In Section 2.1.1, we noted that our study employs a numerical propagator for integration of the equations of motion. During the development of this navigation solution, two numerical propagators were considered:

- an explicit fourth-order Runge-Kutta method using the standard equations (see, e.g., Butcher [97]);
- a built-in MATLAB function (`ode45`) realizing a Runge-Kutta style algorithm developed by Dormand and Prince [98]. It is accurate to the fourth-order with fifth-order evaluations used for error-checking.

The Runge-Kutta framework is appropriate as it has been used in on-orbit filtering applications; these applications include the DIODE algorithm (processing DORIS measurements) and the DIOGENE algorithm (processing GPS measurements) [91]. In our application, we found that `ode45` matched the accuracy of the explicit fourth order method at a fraction of the computational cost. Thus, `ode45` was the integrator of choice.

¹Ideally, this information would be included in the GPS simulator data products, but a bug in the simulator necessitated post-simulation application of the SV antenna pattern.

Appendix D

Solar Radiation Pressure: Derivation and Justification

Solar Radiation Pressure at 1 AU. The Sun constantly emits electromagnetic energy. At a distance of 1 AU from the Sun, the flux of this solar energy has an accepted value of $\Phi(1 \text{ AU}) = 1367 \text{ J}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ ([26], p. 575).

We recall that using the mass-energy equivalence $E = mc^2$, we can define the momentum of electromagnetic energy as $p = mc = \frac{E}{c}$. Using this momentum-energy equivalence, we can use the solar energy flux Φ to define a solar momentum flux Π :

$$\Pi(1 \text{ AU}) = \frac{\Phi(1 \text{ AU})}{c} = \frac{1367 \text{ J}\cdot\text{m}^{-2}\cdot\text{s}^{-1}}{3 \times 10^8 \text{ m}\cdot\text{s}^{-1}} = 4.57 \times 10^{-6} \frac{\text{kg}\cdot\text{m}\cdot\text{s}^{-1}}{\text{m}^2\cdot\text{s}}. \quad (\text{D.1})$$

This momentum flux (momentum transferred per unit time and area) can be equivalently viewed as a force per unit area, a “solar radiation pressure” exerted by the incident solar energy:

$$p_{SR}(1 \text{ AU}) = 4.57 \times 10^{-6} \text{ N}\cdot\text{m}^{-2}. \quad (\text{D.2})$$

For further details, please consult Vallado [26] and Ruiz Delgado [99].

Variation with sun-satellite distance. In the above, we derived the solar radiation pressure at a nominal distance of 1 AU. Both Φ and p_{SR} obey an inverse square law with respect to distance from the Sun. Thus, to obtain p_{SR} at any sun-satellite distance $r_{sat \rightarrow \odot}$, we solve:

$$p_{SR}(r_{sat \rightarrow \odot}) = p_{SR}(1 \text{ AU}) \cdot \left(\frac{1 \text{ AU}}{r_{sat \rightarrow \odot}} \right)^2. \quad (\text{D.3})$$

Variation with solar cycle. Lee et al. [100] found that over Solar Cycle 22, the solar flux at 1 AU varied by at most $1.4 \text{ J}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, approximately 0.1% of the total value. This inaccuracy is negligible when compared to the accuracy of our propagator (see Table 2.5) and can therefore be safely ignored.

Appendix E

GPS Broadcast Errors

The following tables contain the results of an analysis performed by Dr. Anton de Ruiter of Ryerson University. The analysis examined the accuracy of the SV ephemerides and SV clock-corrections broadcast by the GPS constellation. This accuracy can be determined by comparing the broadcast values with highly precise values determined through post-processing. The resultant differences have been used in this thesis to represent typical errors in the broadcast ephemerides and clock-corrections. For further details, including an explanation of the values in the tables, please see the discussions in Section 3.4.3 and Section 3.4.4.

Raw data for the analysis was made available by Dr. Richard Langley at the University of New Brunswick.

Table E.1: SV ephemeris errors (RMS) for the day of 04 Feb 2007.

PRN	Δx , m	Δy , m	Δz , m	PRN	Δx , m	Δy , m	Δz , m
1	-3.414	1.570	2.218	17	2.254	1.391	2.015
2	-1.447	1.095	1.109	18	1.347	0.899	1.065
3	2.663	2.442	-2.699	19	1.279	1.476	1.851
4	-1.505	-1.313	1.669	20	0.882	-1.168	-1.092
5	5.994	5.248	5.258	21	1.969	-1.578	-2.136
6	2.318	-1.979	-2.011	22	-1.162	1.218	1.427
7	1.500	1.316	-1.356	23	-2.079	11.013	-20.265
8	-1.446	-0.846	-1.144	24	0.783	-1.139	1.210
9	-2.187	1.403	2.036	25	1.062	-0.870	-1.525
10	5.062	4.536	-5.683	26	1.285	1.790	-1.095
11	-1.842	-1.259	-1.223	27	4.807	-3.750	-2.726
12	-1.863	1.302	1.001	28	-1.577	1.235	-1.213
13	-1.457	1.728	1.974	29	-5.956	-8.131	-6.041
14	-1.369	-1.006	1.490	30	-5.014	-5.293	4.737
15*	5.609	-3.541	4.343	31	-1.477	1.288	-1.127
16	-1.549	-1.164	0.885	32*	-3.827	2.959	-0.474

*data simulated based on mean and standard deviation of remaining data points.

Table E.2: SV clock-correction errors (RMS) for the day of 04 Feb 2007.

PRN	Δa_{f0} , s	Δa_{f1} , s·s ⁻¹	Δa_{f2} , s·s ⁻²
1	3.76×10^{-9}	-1.60×10^{-12}	2.31×10^{-17}
2	7.28×10^{-10}	-1.76×10^{-13}	1.11×10^{-17}
3	8.09×10^{-9}	-7.61×10^{-13}	9.15×10^{-17}
4	-3.64×10^{-9}	-2.85×10^{-13}	3.31×10^{-17}
5	-9.63×10^{-10}	-1.39×10^{-13}	1.41×10^{-17}
6	-4.79×10^{-9}	-4.13×10^{-14}	-2.17×10^{-18}
7	-2.52×10^{-9}	1.87×10^{-13}	-3.06×10^{-17}
8	-1.09×10^{-9}	-2.21×10^{-13}	1.98×10^{-17}
9	-5.21×10^{-10}	2.22×10^{-14}	2.73×10^{-17}
10	-2.96×10^{-9}	2.25×10^{-13}	1.03×10^{-17}
11	1.02×10^{-8}	-3.93×10^{-14}	1.11×10^{-17}
12	2.45×10^{-9}	1.27×10^{-13}	-8.72×10^{-18}
13	6.90×10^{-9}	-3.51×10^{-14}	9.48×10^{-19}
14	6.60×10^{-9}	-5.60×10^{-14}	-5.97×10^{-18}
15	-2.80×10^{-9}	-1.02×10^{-13}	1.36×10^{-17}
16	5.29×10^{-9}	1.98×10^{-13}	-1.98×10^{-17}
17	1.68×10^{-9}	1.13×10^{-13}	-1.43×10^{-17}
18	4.92×10^{-9}	9.28×10^{-15}	1.20×10^{-18}
19	1.58×10^{-9}	-8.06×10^{-14}	8.11×10^{-18}
20	5.77×10^{-9}	1.24×10^{-13}	-8.25×10^{-18}
21	4.99×10^{-9}	2.20×10^{-13}	-1.80×10^{-17}
22	-2.13×10^{-10}	-1.80×10^{-13}	1.35×10^{-17}
23	1.54×10^{-9}	-7.61×10^{-15}	-7.72×10^{-18}
24	-1.09×10^{-8}	5.99×10^{-13}	-3.07×10^{-17}
25	-1.77×10^{-9}	-3.46×10^{-13}	3.32×10^{-17}
26	-1.10×10^{-9}	1.80×10^{-13}	1.12×10^{-17}
27	1.58×10^{-9}	3.87×10^{-13}	-7.90×10^{-18}
28	1.34×10^{-8}	2.64×10^{-13}	-1.77×10^{-17}
29	-3.97×10^{-9}	-1.17×10^{-13}	1.59×10^{-17}
30	2.85×10^{-10}	1.98×10^{-13}	-1.32×10^{-16}
31	-9.57×10^{-11}	-3.66×10^{-14}	-2.43×10^{-18}
32*	-5.32×10^{-11}	-1.49×10^{-13}	4.60×10^{-17}

*data simulated based on mean and standard deviation of remaining data points.