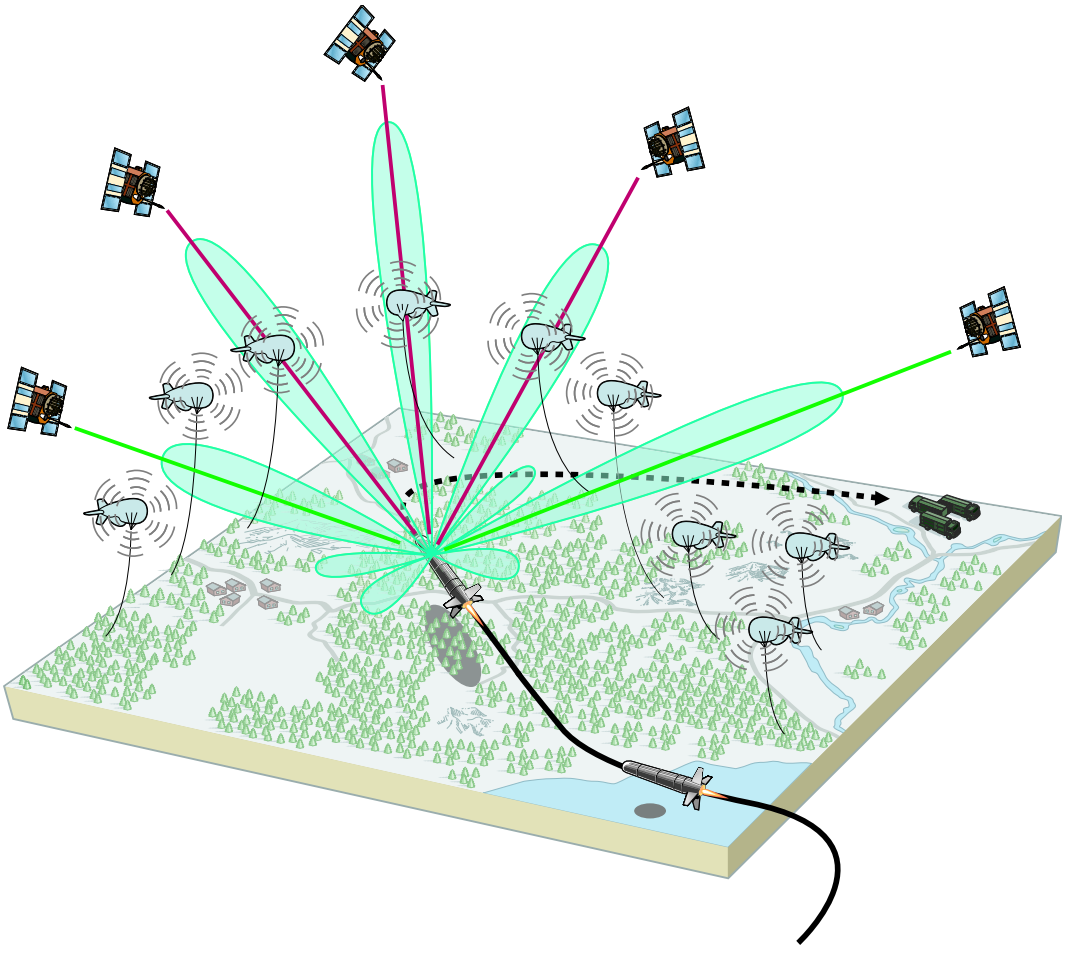
FOI-R--0848--SE
April 2003
1650-1942

Scientific report

Johan Malmström

Robust Navigation with GPS/INS and Adaptive Beamforming



System Technology Division
SE-172 90 STOCKHOLM
Sweden

Swedish Defence Research Agency
System Technology Division
SE-172 90 STOCKHOLM
Sweden

FOI-R--0848--SE
April 2003
1650-1942

Scientific report

Johan Malmström

Robust Navigation with GPS/INS and Adaptive Beamforming

Issuing organization	Report number, ISRN	Report type
Swedish Defence Research Agency System Technology Division SE-172 90 STOCKHOLM Sweden	FOI-R--0848--SE	Scientific report
	Research area code: Electronic Warfare	
	Month year: April 2003	Project no.: E6034
	Customers code: Commissioned Research	
	Sub area code: Electronic Warfare including Electro-magnetic Weapons and Protection	

Author/s (editor/s)	Project manager
Johan Malmström	Bengt Boberg
	Approved by: Monica Dahlén
	Sponsoring agency: Swedish Armed Forces
	Scientifically and technically responsible: John Robinson

Report title

Robust Navigation with GPS/INS and Adaptive Beamforming

Abstract

This report can be divided into two parts, the *Integration of Inertial Navigation System (INS) and GPS* and *Adaptive beamforming for a GPS antenna array*. GPS and INS have complementary properties. Therefore integration results in enhanced performance and robustness compared to each of the two systems alone.

In the first part a tightly coupled navigation filter, based on code pseudorange measurements and an IMU with Litton LN 200 performance has been derived. The filter is used in a comparison of a stand-alone receiver and differential GPS (two receivers). The navigation filter is a complementary extended Kalman estimator, where combinations of GPS pseudoranges and pseudoranges predicted by the INS are used as observations. Systematic errors in the inertial sensors are estimated and compensated for, which gives an improved navigation performance also during periods without GPS aiding. Results are given as simulations where the different integration methods have been tested using partial satellite outages, different flight paths and atmospheric disturbances.

The second part of the report shows how the GPS/INS based navigation system can be made even more robust by using an adaptively beamforming antenna. The received GPS signals are extremely weak and therefore vulnerable to interfering signals. The robustness and performance of the beamforming algorithm LCMV is tested in static simulation scenarios, with multiple jammers and different numbers of array elements. Also, navigation performance is tested for a flying vehicle using the tightly coupled navigation filter in a dynamic scenario with severe jamming from airborne jammers. Using the most effective algorithm, multiple outputs LCMV, increases the equivalent C/N_0 with more than 40 dB, in the dynamic scenario.

Keywords

Navigation, GPS, INS, Sensor integration, Kalman Filter, Jamming, Adaptive beamforming, LCMV.

Further bibliographic information	Language
	English
ISSN: 1650-1942	Pages: 102
Distribution: By sendlist	Price Acc. to pricelist Security classification Unclassified

Utgivare	Rapportnummer, ISRN	Klassificering
Totalförsvarets forskningsinstitut Avdelningen för Systemteknik SE-172 90 STOCKHOLM Sweden	FOI-R--0848--SE	Vetenskaplig rapport
	Forskningsområde: Telekrig	
	Månad, år: April 2003	Projektnummer: E6034
	Verksamhetsgren: Uppdragsfinansierad verksamhet	
	Delområde: Telekrigföring med EM-vapen och skydd	

Författare/redaktör	
Johan Malmström	Projektledare: Bengt Boberg
	Godkänd av: Monica Dahlén
	Uppdragsgivare/kundbeteckning: Försvarsmakten
	Tekniskt och/eller vetenskapligt ansvarig: John Robinson

Rapportens titel

Robust navigering med GPS/tröghetsnavigering och adaptiv lobformning

Sammanfattning

Denna rapport består av två delar, *Integrering av tröghetsnavigering och GPS* samt *Adaptiv lobformning för GPS med en gruppantenn*. GPS och tröghetsnavigering har komplementära egenskaper. Därför ger en integrering bättre navigeringsprestanda och robusthet än systemen var för sig.

I den första delen av rapporten presenteras ett tätt kopplat navigeringsfilter baserat på kod-pseudoavstånds-mätningar och signaler från tröghetssensorer liknande Litton LN 200. Filtret används vid en jämförelse av fallen med en ensam GPS-mottagare och en differentiell GPS (två mottagare). Navigeringsfiltret är en Kalmanestimator (EKF) vilken som observationer använder kombinationer av GPS-pseudoavstånd och pseudoavstånd predikterade av tröghetsnavigeringen. Systematiska fel i tröghetssensorerna estimeras och kompenseras för, vilket ger en förbättrad navigeringspresanda också under perioder utan GPS-stöd. Resultat presenteras i form av simuleringar där olika integreringsmetoder testas med satellitbortfall, olika flygbanor och atmosfärspåverkan av signalerna.

De mottagna GPS-signalerna är extremt svaga och därför känsliga för störningar. Den andra delen av rapporten visar hur störfastheten hos det GPS/INS-baserade navigeringssystemet kan förbättras ytterligare med hjälp av adaptiv lobformning. Störfasthet och prestanda för lobformningsalgoritmen LCMV testas i statiska simuleringsscenarier för flera störare och olika antal antennelement. Navigeringsprestanda undersöks för en flygande farkost i ett dynamiskt scenario med svår störning från luftburna störare. Med den effektivaste lobformningsalgoritmen, parallella LCMV, ökar det ekvivalenta C/N_0 med över 40 dB i det dynamiska scenariot.

Nyckelord

Navigering, GPS, INS, Sensorintegrering, Kalmanfilter, Störning, Adaptiv lobformning, LCMV.

Övriga bibliografiska uppgifter	Språk
	Engelska

ISSN	Antal sidor
1650-1942	102

Distribution	
Enligt missiv	Pris: Enligt prislista Sekretess: Öppen

Preface

This report is the result of my Master's project. It is the last part on my way to a M.Sc. degree in Electrical Engineering at the Royal Institute of Technology (KTH) in Stockholm, Sweden.

The work was carried out at the Swedish Defence Research Agency (FOI) at the department for Autonomous Systems in Ursvik, Stockholm. I would like to thank my supervisor at FOI, Bengt Boberg, for his help and support in this work, and my supervisor and examiner at the department of Signals, Sensors & Systems (S3) at KTH, Magnus Jansson.

I would also like to thank Fredrik Berefelt and Sven-Lennart Wirkander for their interest in my work and spending their time helping me. I also appreciate all help I got from Lars Pääjärvi and Fredrik Eklöf at FOI in Linköping, they have been a great resource in the topics of GPS and adaptive beamforming.

At last I would like to thank the Swedish Radio Navigation Board (RNN) for awarding me a scholarship for best Master's project 2003.

Parts of the results in this report were presented by FOI at the ION GPS/GNSS conference in Portland USA, September 9–12, 2003.

Johan Malmström
Stockholm, October 2003

Contents

1. Introduction

Knowing your position is, in many situations, very important. It is so today and has been so for a long time. Today many navigation systems are based on Global Navigation Satellite Systems (GNSS), of which the US system Global Positioning System (GPS) is the most widely used. Because of very low signal power levels from the satellites, this system is vulnerable to interfering signals. Increased robustness can be achieved by integrating GPS with another navigation system. An attractive solution is integration with an Inertial Navigation System (INS). An INS is a self contained system. The position of a vehicle is calculated using measurements of accelerations and rotations in three dimensions, which are all quantities that can be measured on the vehicle without external aiding or data. This makes an INS extremely insensitive to external interferences, i.e. jamming.

Another way to protect the GPS receiver from interfering signals is to use a more sophisticated antenna. With an antenna array, consisting of several antenna elements, it is possible to direct a beam with higher gain in the directions of the satellites. This can give a significantly increased signal-to-noise ratio (SNR) that in turn increases the availability of the GPS.

1.1 Objective

The objective of this thesis is to create a simulation environment for a robust navigation system. The environment should contain a centralized integration of an INS and a GPS receiver, using a tightly coupled complementary Kalman filter.

To increase the availability during periods when other signals are interfering the GPS signals, simulations with an adaptive beamforming antenna should be performed.

The work is divided into two parts, *Integration of INS/GPS* and *Beamforming for GPS*.

1.2 Problem and Methods

1.2.1 Integration of INS and GPS INS and GPS have several complementary properties, e.g. they have small errors for short and long times, respectively. The GPS system is based on distance measurements and can provide position and time estimates with bounded errors. The sample rate is not very high, typically a few Hertz, which is often too low for control purposes. An INS, on the other hand, can provide measurements at higher sampling rates, but as noisy sensor data are integrated its error without bound. The fact that an INS and a GPS have these complementary properties can be used advantageously, if the two systems are integrated. The integrated system can therefore have better navigation performance and robustness than each of the systems alone. This behaviour is depicted in Figure 1.1.

The integration of raw IMU sensor signals and GPS pseudoranges is performed by means of a complementary Kalman filter. The filter structure used is depicted in Figure 1.2.

1.2.2 Beamforming for GPS Beamforming means that the gain of an antenna is concentrated to a certain direction. In this thesis an *antenna array* is used, which consists of a number of antenna elements. By manipulating the signals from each element the *reception pattern*, which describes the antenna gain for all directions, can be controlled. When using a beamforming antenna for GPS the beams must be controlled in three dimensions since the orbits of the GPS satellites can cover the whole sky. An example of a three-dimensional reception pattern is shown in Figure 1.3.

1.3 Previous Work

The area of inertial navigation and its performance has been well known for a long time. The developments in INS performance are mainly due to better performance in the sensors.The development of satellite based navigation systems in the 1990's, e.g. GPS, and the ability to integrate the two systems, gave a rebirth to the INS research. Yet today, this is an area that is not considered fully investigated.

The second main area of this thesis, *adaptive beamforming*, is a relatively new topic. The theories and many algorithms were developed in the late 1980's, but it is not before the last years, with the increased computational power of computers, that the algorithms have been able to be used in real applications.

At the Swedish Defence Research Agency (FOI) some relevant research has been made within the field of GPS and INS integration. The Kalman estimator and a particle estimator filter have been compared in a navigation application [1] and the dynamics of an INS system have been modelled and implemented in FOI-NAV MATLAB toolbox. Another project at FOI has developed a MATLAB toolbox for adaptive beamforming antennas in communication systems [2].

1.4 Outline of Thesis

The thesis is divided into several chapters. Chapters 2–4 mainly present background information and theories.

Chapter 2 Inertial Navigation describes the principles of inertial navigation. Some mathematical terms for transformation between different systems are introduced. The well-known navigation equations and their error dynamics are derived.

Chapter 3 Satellite Navigation explains how the Global Positioning System (GPS) works. The performance of a GPS receiver and errors affecting the accuracy are

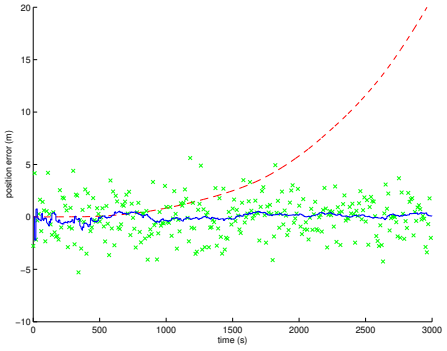


Figure 1.1: Comparison of error performance between integrated and non-integrated systems. The integrated system (solid line) performs better than both an unaided INS (dashed line) and a stand alone GPS (marked with ×). The picture is generated with the simulation environment developed in this project.

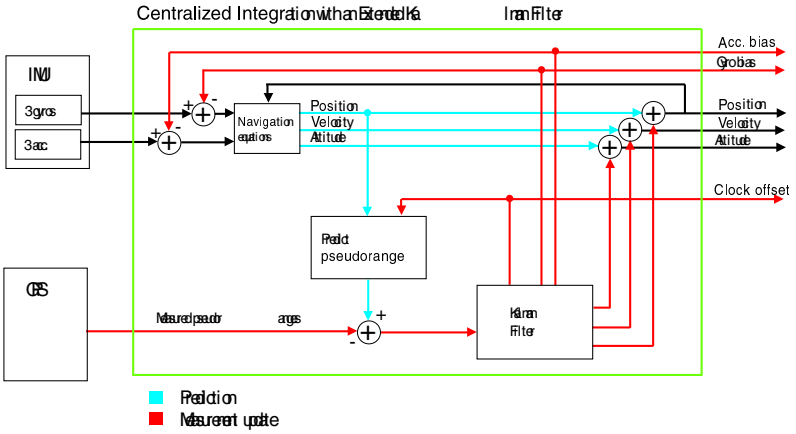


Figure 1.2: The tight integration of INS and GPS is done by means of a complementary Extended Kalman filter. The inputs to the filter are raw data from the INS and GPS, i.e. gyro and accelerometer data and pseudoranges to satellites.

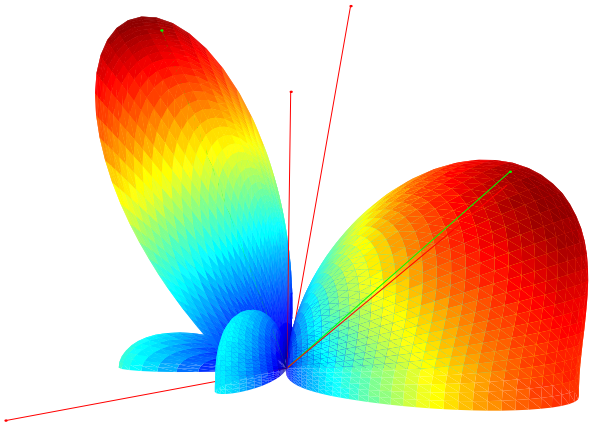


Figure 1.3: The signal-to-noise ratio can be significantly increased by using an adaptively beamforming antenna. A beam with higher gain can be directed towards a satellite in the same time as interfering signals are suppressed. In the example above, two main beams are formed toward two satellites and the noise signals (red lines) are heavily suppressed.

discussed. The chapter is written to give the reader a general knowledge about GPS. In [3] a much more detailed description of GPS is given.

Chapter 4 Adaptive Beamforming for GPS Some theory for adaptive beamforming is presented. Some beamforming algorithms proposed in the literature and articles are described.

Chapters 5–6 are more applied to this specific problem.

Chapter 5 Linearization and Discretization of Nonlinear Systems In this chapter the non-linear error dynamics of the navigation equations are linearized and discretized to fit to the general state-space form.

Chapter 6 Implementation of Simulation Environment describes how the two systems, the INS and the GPS, were modelled, and how the integration filter was designed.

The three last chapters present results in form of simulations and summarizes the work.

Chapter 7 Simulation Results for GPS/INS Integration Simulations of the system performance have been made in MATLAB. The integration filter is tested in different situations.

Chapter 8 Simulations of Jamming Suppression using Adaptive Beamforming The performance when using an adaptive beamforming antenna with the integrated INS/GPS is tested.

Chapter 9 Conclusions summarizes the work. Some suggestions for future work is given.

2. Inertial Navigation

The principles of inertial navigation are actually rather simple. For short, it is about measuring relative transportation from a given initial position and with a given initial velocity. From the start position a measured acceleration is integrated twice with respect to time, in order to get the relative transportation. To know in which direction the transportation is taking place, it is also necessary to measure the change in attitude.

Since the problem is solved by integration, small measurement errors are integrated, which has the effect that the calculated position "drifts away" from the true position. Especially, an INS is sensitive to biases in the sensors. High-frequency noise is attenuated in the integration, which is a low-pass process.

The drift is one of the major drawbacks with inertial navigation. There are high-precision sensors with very small biases and noise, but nevertheless, there will be a drift error. An unaided INS has an unbounded long-term error.

It is worth noting that the integration process only requires information that can be measured by the user himself, a so called self-contained system. The advantage of having a self-contained system, the INS, is that it is a very robust way of navigating. It is practically impossible to disturb an INS except by physically destroying the inertial sensors.

More details about inertial navigation can be found in [4], [5] and [1].

2.1 The Principle of Inertial Navigation

The foundation of inertial navigation is *Newton's Laws of motion*. Newton's first law says that in an inertial frame, a body at rest or in uniform motion will remain so in absence of applied forces.

The second law states that the position $\mathbf{x}$ of a body with mass m affected by a force $\mathbf{f}_g$, where gravity is excluded, can be described with the relation

$$\mathbf{f}_g = \frac{d}{dt}(m\dot{\mathbf{x}}). \tag{2.1}$$

Neglecting relativistic effects and assuming a constant mass gives

$$\mathbf{f}_g = m\ddot{\mathbf{x}}. \tag{2.2}$$

This relation is valid in an inertial frame. However, a philosophical problem with defining an inertial frame remains. An Earth-fixed frame is not inertial since Earth is rotating. A frame fixed in space, where Earth is rotating is not really inertial either, since this frame is affected by the gravity field from the sun. Also our solar system is rotating around a point in the expanding universe. The only inertial frame, at least in a local sense, where the classical mechanics hold, is a frame under free fall.

For practical purposes, the frame attached to the center of Earth and not rotating with respect to distant stars (Figure 2.1), can be considered inertial [4]. This requires

that compensation is done, at least for the gravity from Earth. Otherwise a body would not stay in rest if no force is applied, as stated by Newton's first law. With the gravitational field $\mathbf{g}$, affecting a mass m, equation (2.2) becomes

$$\mathbf{f} + m\mathbf{g} = m\ddot{\mathbf{x}} \tag{2.3}$$

or, after division with the mass m,

$$\ddot{\mathbf{x}} = \frac{\mathbf{f}}{m} + \mathbf{g}. \tag{2.4}$$

There are methods for compensating not only for the gravity caused by Earth in the term $\mathbf{g}$, but also for the gravity from the moon and sun and other celestial bodies. In this thesis we are satisfied with considering the gravity from Earth.

The term $\frac{\mathbf{f}}{m}$ in (2.4) is recognized as an acceleration, and will be denoted as $\mathbf{a}$. The acceleration can be measured with an accelerometer. With knowledge of $\mathbf{a}$ and the gravity $\mathbf{g}$, the differential equation (2.4) could be integrated twice, to get the position $\mathbf{x}$. However, we also need to know the orientation of the vehicle, to know in what direction the integration of (2.4) should take place. The orientation can be calculated from angular rates, measured by a gyro. Integrating the gyro output yields the relative change in orientation. Therefore the current orientation can be estimated, if the initial orientation is known.

A set of sensors used for inertial navigation, an *Inertial Measurement Unit* (IMU), usually consists of three gyros and three accelerometers, put together in two mutually orthogonal sets. Such a set can measure acceleration and angular rates in three dimensions. Later, we will se that this is sufficient for solving (2.4) for $\mathbf{x}$. More about performance and different types of IMUs can be found in [4].

To continue the derivation of the equations needed for inertial navigation, we first have to define different frames and transformations between them.

2.2 Coordinate Frames

Different coordinate frames are used depending on the application. It is important to sort out in which frame each quantity is measured. In all frames, positions will be defined in right-handed coordinate systems with orthogonal axes. A quantity x, measured in the arbitrary a-frame, will be denoted x^a (if it is necessary to point it out).

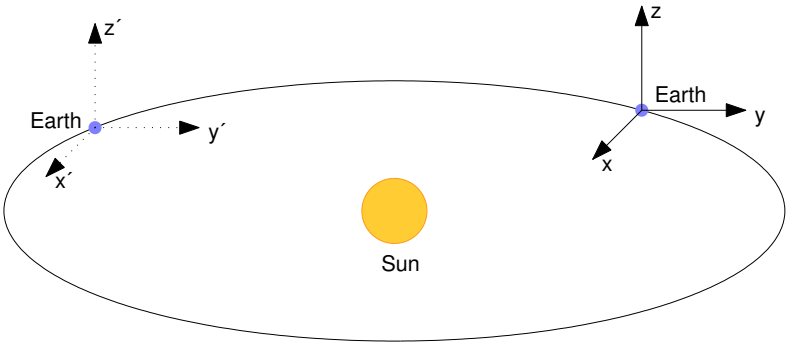


Figure 2.1: Earth-Centered Inertial frame (ECI or i-frame). The ECI frame is in free fall in the gravity field of the sun, but not rotating with respect to the distant fixed stars.

Earth Centered Systems Earth centered systems have, as the name indicates, their origins in the center of Earth. These frames are preferred if the equations should be valid irrespective of position on Earth.

ECI Earth-Centered Inertial frame (i-frame). It is the changes relative to this frame that are measured by the IMU.

The origin of the frame is located at the center of earth. The 1- and 2-axis span the Equator plane and the 3-axis coincides with the rotation axis of Earth and pointing towards the north pole. Figure 2.1 illustrates the ECI frame.

ECEF Earth-Centered Earth-Fixed frame (e-frame). The ECEF frame is similar to the ECI frame, but this frame rotates with the same angular rate ω_e as Earth[1]. Thanks to the rotation, a fixed point on Earth also has fixed coordinates in this frame.

Coordinates in e-frame can also be expressed as polar coordinates (longitude, latitude and height). These coordinates are referred to as *Geodetic coordinates*, but will not be used in this thesis.

Local Systems When navigating in a local area where Earth can be considered flat it is convenient to give the coordinates relative to a fixed reference point. To give a coordinate in a local system the coordinates of the reference point must be defined.

When a local system is used in this thesis, the coordinates will be given as Cartesian coordinates (East, North, Up).

Vehicle Centered Systems These systems are moving with the vehicle, so that the vehicle is always centered in the origin of the frames.

Body Fixed Frame The Body Fixed Frame (b-frame) is defined with the 1-axis in the front direction of the vehicle, the 2-axis pointing to the right and the 3-axis through the floor.

The IMU can either be gimballed or fixed to this frame. Here, the IMU is considered to be strapped to the vehicle, a so called *strapped down INS*. The IMU can also be arranged in a *gimballed system*, so that the IMU can turn relative to the b-frame. In that case the IMU is aligned with another well-defined frame, i.e. the local frame.

2.3 Coordinate Transformations

In section 2.2 the different frames that will be used was described. It is important to be able to transform coordinates expressed in one frame to another. If the directions of the coordinate axes differ, the coordinates have to be rotated.

2.3.1 Direction Cosine Matrix The direction cosine matrix (DCM) is used to transform coordinates between two concentric frames. Direction cosine matrices are important to be able to understand the derivation of the navigation equations which follows.

A transformation of coordinates $\mathbf{x}^t$ in the t-frame to coordinates $\mathbf{x}^s$ in the s-frame can be done as

$$\mathbf{x}^s = \mathbf{C}_t^s \mathbf{x}^t \tag{2.5}$$

[1] $\omega_e = 7.292115 \cdot 10^{-5}$ rad/s given by WGS-84 [6]

where the direction cosine matrix is

$$\mathbf{C}_t^s = \begin{bmatrix} c_{1,1} & c_{1,2} & c_{1,3} \\ c_{2,1} & c_{2,2} & c_{2,3} \\ c_{3,1} & c_{3,2} & c_{3,3} \end{bmatrix}. \tag{2.6}$$

The columns are mutually orthogonal and so are the rows. This implies that $\mathbf{C}_t^s$ is an orthogonal matrix with the following properties:

$$\mathbf{C}_t^s(\mathbf{C}_t^s)^T = I \quad \Rightarrow \quad \mathbf{C}_s^t \equiv (\mathbf{C}_t^s)^{-1} = (\mathbf{C}_t^s)^T \tag{2.7}$$

A theorem by Euler, referred to in [4], says that any sequence of rotations can be represented by one single rotation about a single axis. This fact can be used to perform any coordinate transforms by a number of rotations around the coordinate axes. In three dimensions, the rotation is split up into three rotations, one around each coordinate axis. Together, the rotations give the same result as one rotation around the axis as given by Euler's theorem. For small angles the order of rotation does not matter.

One important property that can be expressed in terms of rotation around the coordinate axes is the time derivative of a transition matrix $\mathbf{C}_t^s$. In [4] it is shown that $\dot{\mathbf{C}}_t^s$ can be expressed as

$$\dot{\mathbf{C}}_t^s = \mathbf{C}_t^s \, \mathbf{\Omega}_{st}^t. \tag{2.8}$$

Here, the factor $\mathbf{\Omega}_{st}^t$ is a *skew-symmetric matrix* which is built up by $\boldsymbol{\omega}_{st}^t = [\omega_1, \, \omega_2, \, \omega_3]$. The subscripts in $\boldsymbol{\omega}_{st}^t$ tells that it is the rotation of the t-frame relative to the s-frame that is measured, and the superscript tells that the rotation is measured with coordinates from the t-frame.

$$\mathbf{\Omega}_{st}^t = [\boldsymbol{\omega}_{st}^t \times] = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}. \tag{2.9}$$

The notation $[\boldsymbol{\omega}_{st}^t \times]$ is sometimes used for $\mathbf{\Omega}_{st}^t$ and comes from the close relation with the vector cross product. Time-derivatives of small angles, like $\boldsymbol{\omega}_{st}^t$ above that describe rotation rates, are usually called *angular rates*.

2.3.2 Quaternions Multiplication by a complex number $s = \|s\| \, e^{j \arg(s)}$ with the constraint $\big\|s\big\| = 1$, can be seen as a rotation of an arbitrary complex number $c = \|c\| \, e^{j \arg(c)}$. Since

$$\begin{aligned} s \cdot c &= \|s\| \, e^{j \arg(s)} \cdot \|c\| \, e^{j \arg(c)} \\ &= 1 \cdot e^{j \arg(s)} \cdot \|c\| \, e^{j \arg(c)} \\ &= \|c\| \, e^{j(\arg(c) + \arg(s))} \end{aligned} \tag{2.10}$$

the complex number c is rotated around the origin with the angle $\arg(s)$.

A quaternion is an extension of the complex numbers that perform the rotation described in (2.10) in three dimensions.

$$q = j \cdot a + k \cdot b + l \cdot c + d \tag{2.11}$$

where j, k and l are mutually orthogonal imaginary units that span a right-hand coordinate system. There is still a unit constraint so that $\|q\| = 1$. The constraint reduces the number of degrees of freedom from 4 to 3, in the same way as the unit constraint reduces the unit complex number from 2 to 1 degree of freedom.

Quaternions are used in the integration process of the navigation equations that is described later in this chapter. However, the quaternions are not crucial for understanding the further description. More about quaternions can be found in [4] or [7].

2.4 Navigation Equations

The navigation equations are a set of differential equations describing how position, velocity and attitude are changed depending on acceleration and angular rotations. In this thesis the navigation equations will be expressed with coordinates in the e-frame (see Section 2.2).

Applying (2.4) gives the following equation for the position coordinates $\mathbf{x}^i$ in the inertial frame (i-frame):

$$\ddot{\mathbf{x}}^i = \mathbf{a}^i + \mathbf{g}^i. \tag{2.12}$$

We want to develop a corresponding expression valid in the e-frame. A coordinate $\mathbf{x}^e$ can be expressed in the coordinates $\mathbf{x}^i$ by means of the transition matrix $\mathbf{C}_e^i$ according to

$$\mathbf{x}^i = \mathbf{C}_e^i \, \mathbf{x}^e \tag{2.13}$$

The time derivative of the transition matrix $\mathbf{C}_e^i$ can be calculated if the general equation (2.8) is applied to $\mathbf{C}_e^i$:

$$\dot{\mathbf{C}}_e^i = \mathbf{C}_e^i \, \mathbf{\Omega}_{ie}^e \tag{2.14}$$

where $\mathbf{\Omega}_{ie}^e$ is a skew-symmetric matrix with elements given by the angular rates $\boldsymbol{\omega}_{ie}^e = [\omega_1, \omega_2, \omega_3]$ between the i-frame and e-frame:

$$\mathbf{\Omega}_{ie}^e = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}. \tag{2.15}$$

The only non-zero components in $\mathbf{\Omega}_{ie}^e$ is ω_3, since the e-frame is defined to rotate about the 3-axis. This component is denoted ω_e.

We further need the second time derivative of $\mathbf{C}_e^i$, which is

$$\ddot{\mathbf{C}}_e^i = \mathbf{C}_e^i \, \dot{\mathbf{\Omega}}_{ie}^e + \mathbf{C}_e^i \, \mathbf{\Omega}_{ie}^e \, \mathbf{\Omega}_{ie}^e, \tag{2.16}$$

where (2.14) has been used twice. It is possible to simplify this expression by noting that the Earth rotation $\mathbf{\Omega}_{ie}^e$ is constant, and therefore $\dot{\mathbf{\Omega}}_{ie}^e$ is zero, which gives

$$\ddot{\mathbf{C}}_e^i = \mathbf{C}_e^i \, \mathbf{\Omega}_{ie}^e \, \mathbf{\Omega}_{ie}^e. \tag{2.17}$$

Now, differentiating (2.13) and inserting the expressions for the first and second derivative of $\mathbf{C}_e^i$ from (2.14) and (2.17) yields

$$\begin{aligned} \ddot{\mathbf{x}}^i &= \ddot{\mathbf{C}}_e^i \, \mathbf{x}^e + 2 \, \dot{\mathbf{C}}_e^i \, \dot{\mathbf{x}}^e + \mathbf{C}_e^i \, \ddot{\mathbf{x}}^e \\ &= \mathbf{C}_e^i \, \mathbf{\Omega}_{ie}^e \, \mathbf{\Omega}_{ie}^e \, \mathbf{x}^e + 2 \, \mathbf{C}_e^i \, \mathbf{\Omega}_{ie}^e \, \dot{\mathbf{x}}^e + \mathbf{C}_e^i \, \ddot{\mathbf{x}}^e. \end{aligned} \tag{2.18}$$

Multiplying both sides with $\mathbf{C}_i^e = (\mathbf{C}_e^i)^{-1}$, solving for $\ddot{\mathbf{x}}^e$ and using (2.12) gives

$$\ddot{\mathbf{x}}^e = -2 \, \mathbf{\Omega}_{ie}^e \, \dot{\mathbf{x}}^e - \mathbf{\Omega}_{ie}^e \mathbf{\Omega}_{ie}^a \, \mathbf{x}^e + \mathbf{a}^e + \mathbf{g}^e \tag{2.19}$$

where $\mathbf{g}^e = \mathbf{C}_i^e \, \mathbf{g}^i$ are the components of the gravity expressed in the e-frame. The quantity $\mathbf{a}^e$ is the sensed acceleration in the e-frame. The sensed acceleration in e-frame is given by

$$\mathbf{a}^e = \mathbf{C}_b^e \, \mathbf{a}^b, \tag{2.20}$$

where $\mathbf{a}^b$ are the components in the b-frame of the sensed acceleration. The transition matrix $\mathbf{C}_b^e$ from b-frame to e-frame in (2.20), that describes the attitude of the vehicle, is determined from the gyro outputs. The dynamics of the attitude is described by the general equation (2.8) which, applied to the e- and b-frame, gives

$$\dot{\mathbf{C}}_b^e = \mathbf{C}_b^e \, \mathbf{\Omega}_{eb}^b. \tag{2.21}$$

If the angular rates between the e-frame and the b-frame are denoted $\boldsymbol{\omega}_{eb}^b = (\omega_1\ \omega_2\ \omega_3)^T$, the skew-symmetric matrix $\boldsymbol{\Omega}_{eb}^b$ is, according to (2.9),

$$\boldsymbol{\Omega}_{eb}^b = [\boldsymbol{\omega}_{eb}^b \times] = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}. \tag{2.22}$$

But the gyros sense the angular rates relative to the inertial frame, $\boldsymbol{\omega}_{ib}^b$:

$$\boldsymbol{\omega}_{ib}^b = \boldsymbol{\omega}_{ie}^b + \boldsymbol{\omega}_{eb}^b. \tag{2.23}$$

We have to compensate for the angular rates $\boldsymbol{\omega}_{ie}^b$ that is caused by Earth rotation in the i-frame:

$$\begin{aligned} \boldsymbol{\omega}_{eb}^b &= \boldsymbol{\omega}_{ib}^b - \boldsymbol{\omega}_{ie}^b \\ &= \boldsymbol{\omega}_{ib}^b - \mathbf{C}_e^b \boldsymbol{\omega}_{ie}^e. \end{aligned} \tag{2.24}$$

The three-dimensional second-order differential equation (2.19) can be transformed to a system with six first-order differential equations by introducing the velocity vector $\dot{\mathbf{x}}^e$ as state variables. Together with (2.21) and (2.24), a system of nine first-order differential equations, that is sufficient for describing the navigation problem can be formed:

$$\begin{aligned} \frac{d}{dt}\mathbf{C}_b^e &= \mathbf{C}_b^e\,\boldsymbol{\Omega}_{eb}^b \\ \frac{d}{dt}\dot{\mathbf{x}}^e &= -2\,\boldsymbol{\Omega}_{ie}^e\,\dot{\mathbf{x}}^e - \boldsymbol{\Omega}_{ie}^e\boldsymbol{\Omega}_{ie}^e\,\mathbf{x}^e + \mathbf{C}_b^e\,\mathbf{a}^b + \mathbf{g}^e \\ \frac{d}{dt}\mathbf{x}^e &= \dot{\mathbf{x}}^e. \end{aligned} \tag{2.25}$$

The system (2.25) is the navigation equations in e-frame. It is a system of nine differential equations. The only non-linearity in the system is the gravity term, which depends on the position in a non-linear way (see Section 2.4.1). Measurements from accelerometers and gyros are included in $\mathbf{a}^b$ and $\boldsymbol{\Omega}_{eb}^b$, respectively.

In this thesis the ECEF frame (e-frame) will be used for the navigation equations. The reason for this is not only that the navigation equations have a simple form in this frame, but also that GPS expresses coordinates in this frame. Obtaining the navigation equations in other frames is not as straight-forward as for the e-frame. In [4] and [1] the navigation equations can be found for other frames, e.g. the popular, but complicated navigation-frame (n-frame).

2.4.1 Gravity Model Earth will be assumed to be homogeneously dense spherical. With this assumption the gravity will always be directed towards the center of the Earth. An approximation of the gravity according to [4] is

$$\mathbf{g}^e \approx \frac{-kM}{|\mathbf{x}^e|^3}\,\mathbf{x}^\mathbf{e} = \frac{-kM}{r^3}\,\mathbf{x}^\mathbf{e}. \tag{2.26}$$

Here $\mathbf{x}^e = (x_1, x_2, x_3)^T$ are the coordinates in e-frame and, hence, $r = |\mathbf{x}^e|$ is the distance from the center of Earth to the point $\mathbf{x}^e$, where the gravity is evaluated. The constant term

$$kM \approx 3.986 \cdot 10^{14}\ \mathrm{m^3 s^{-2}}$$

is the gravitational constant k times the mass of Earth M.

The derivative of the gravity, the *gravity gradient* $\boldsymbol{\Gamma}^e$ derived from the approximation

in (2.26) is

$$\begin{aligned}\mathbf{\Gamma}^e \equiv \frac{\partial \mathbf{g}^e}{\partial \mathbf{x}^e} &= -kM\left(\frac{\mathbf{I}_3}{r^3} - 3\frac{\mathbf{x}^e(\mathbf{x}^e)^T}{r^5}\right) \\ &= -kM\left(\left[\begin{array}{ccc}1&0&0\\0&1&0\\0&0&1\end{array}\right] r^{-3} - 3\left[\begin{array}{ccc}x_1^2&x_1x_2&x_1x_3\\x_2x_1&x_2^2&x_2x_3\\x_3x_1&x_3x_2&x_3^2\end{array}\right] r^{-5}\right).\end{aligned} \tag{2.27}$$

2.5 Error Dynamics

Taking the differential of system (2.25) shows the dynamics of small errors in the variables (denoted by the prefix δ):

$$\begin{aligned}\frac{d}{dt}\,\delta\mathbf{C}_b^e &= \delta\mathbf{C}_b^e\,\mathbf{\Omega}_{eb}^b + \mathbf{C}_b^e\,\delta\mathbf{\Omega}_{eb}^b \\ \frac{d}{dt}\,\delta\dot{\mathbf{x}}^e &= -2\,\delta\mathbf{\Omega}_{ie}^e\,\dot{\mathbf{x}}^e - 2\,\mathbf{\Omega}_{ie}^e\,\delta\dot{\mathbf{x}}^e - (\,\delta\mathbf{\Omega}_{ie}^e\mathbf{\Omega}_{ie}^e + \mathbf{\Omega}_{ie}^e\delta\mathbf{\Omega}_{ie}^e)\,\mathbf{x}^e \\ &\quad - \mathbf{\Omega}_{ie}^e\mathbf{\Omega}_{ie}^e\,\delta\mathbf{x}^e + \delta\mathbf{C}_b^e\,\mathbf{a}^b + \mathbf{C}_b^e\,\delta\mathbf{a}^b + \delta\mathbf{g}^e + \frac{\partial\mathbf{g}^e}{\partial\mathbf{x}^e}\delta\mathbf{x}^e \\ \frac{d}{dt}\,\delta\mathbf{x}^e &= \delta\dot{\mathbf{x}}^e.\end{aligned} \tag{2.28}$$

The last term in the second equation is the gravity gradient defined in (2.27) and comes from the fact that the gravity $\mathbf{g}^e$ depends on the position $\mathbf{x}^e$. The term $\delta\mathbf{g}^e$ is the gravity disturbance vector. It reflects the gravity deviation from the used gravity model. We assume that the gravity model in Section 2.4.1 describes the gravity with a sufficient accuracy and, hence, the gravity disturbance vector $\delta\mathbf{g}^e$ is set to zero.

The term $\delta\mathbf{C}_b^e$ represents a misalignment in the transition matrix $\mathbf{C}_b^e$, that transforms coordinates from the b-frame to the e-frame. Under the assumption that this misalignment is small, $\delta\mathbf{C}_b^e$ can be written as a skew-symmetric matrix. An alignment error represented by the small angles $\mathbf{\Psi}^e = (\psi_1^e, \psi_2^e, \psi_3^e)^T$ can then, according to [4], be written as

$$\delta\mathbf{C}_b^e = -[\mathbf{\Psi}^e\times]\mathbf{C}_b^e, \tag{2.29}$$

where $[\mathbf{\Psi}^e\times]$ is a skew-symmetric matrix, with elements $(\psi_1^e, \psi_2^e, \psi_3^e)^T$, like the one defined in (2.9). Further, $\mathbf{\Omega}_{ie}^e$ is a known constant and therefore $\delta\mathbf{\Omega}_{ie}^e = 0$.

The error dynamics of the system in (2.25) can now be written as

$$\begin{aligned}\frac{d}{dt}\,\delta\mathbf{C}_b^e &= \delta\mathbf{C}_b^e\,\mathbf{\Omega}_{eb}^b + \mathbf{C}_b^e\,\delta\mathbf{\Omega}_{eb}^b \\ \frac{d}{dt}\,\delta\dot{\mathbf{x}}^e &= -2\,\mathbf{\Omega}_{ie}^e\,\delta\dot{\mathbf{x}}^e - (\mathbf{\Omega}_{ie}^e\mathbf{\Omega}_{ie}^e - \mathbf{\Gamma}^e)\,\delta\mathbf{x}^e + \delta\mathbf{C}_b^e\,\mathbf{a}^b + \mathbf{C}_b^e\,\delta\mathbf{a}^b \\ \frac{d}{dt}\,\delta\mathbf{x}^e &= \delta\dot{\mathbf{x}}^e.\end{aligned} \tag{2.30}$$

with $\mathbf{\Gamma}^e$ as a short notation for the gravity gradient, defined in (2.27). The system can be written in matrix form if the DCM $\delta\mathbf{C}_b^e$ is rewritten in terms of the angles $\mathbf{\Psi}^e$ and a skew-symmetric matrix instead:

$$\frac{d}{dt}\left[\begin{array}{c}\mathbf{\Psi}^e\\\delta\dot{\mathbf{x}}^e\\\delta\mathbf{x}^e\end{array}\right] = \left[\begin{array}{ccc}-\mathbf{\Omega}_{ie}^e&\mathbf{0}_3&\mathbf{0}_3\\[\mathbf{a}^e\times]&-2\,\mathbf{\Omega}_{ie}^e&\mathbf{\Upsilon}\\\mathbf{0}_3&\mathbf{I}_3&\mathbf{0}_3\end{array}\right]\left[\begin{array}{c}\mathbf{\Psi}^e\\\delta\dot{\mathbf{x}}^e\\\delta\mathbf{x}^e\end{array}\right] + \left[\begin{array}{cc}-\mathbf{C}_b^e&\mathbf{0}_3\\\mathbf{0}_3&\mathbf{C}_b^e\\\mathbf{0}_3&\mathbf{0}_3\end{array}\right]\left[\begin{array}{c}\delta\boldsymbol{\omega}_{ib}^b\\\delta\mathbf{a}^b\end{array}\right] \tag{2.31}$$

where $\mathbf{\Upsilon} = -(\mathbf{\Omega}_{ie}^e\mathbf{\Omega}_{ie}^e - \mathbf{\Gamma}^e)$. The terms $\delta\boldsymbol{\omega}_{ib}^b$ and $\delta\mathbf{a}^b$ are errors in the accelerometers and gyros, respectively.

3. Satellite Navigation

In this chapter the principles of Global Navigation Satellite Systems (GNSS) will be described. Since GPS (Global Positioning System) is the by far most commonly used GNSS, the text will be concentrated on this system.

Today two systems for satellite navigation exist, the well-known GPS and a Russian system named GLONASS (GLobal Orbit NAvigation Satellite System). The two systems have many similarities. GLONASS is currently (November 2002) not fully operational, since only about 10 satellites are in their orbits. Normally both GPS and GLONASS have satellite constellations of 24 satellites orbiting earth.

There is also a European system, *Galileo*, under construction. It is planned to be fully operational in year 2008. A first test satellite will be launched in late 2004.

3.1 The Global Positioning System

The GPS system can be divided into three parts or segments: the two system segments, the satellites and ground controls and the user segment, which is all receivers. There are at least 24 satellites orbiting Earth about 20200 km (roughly 3 times the radius of Earth) above the surface. The satellites are tracked by the ground control, which periodically uploads orbit and clock correction data to the satellites. The receivers can calculate their 3–dimensional position by measuring the distances to the satellites.

GPS was designed as a military system by the USA, but is free for civilian users too. Civilian users are restricted to a *Standard Positioning Service* (SPS). The military part of the system, which only the military of the USA and users authorized by them have access to, is referred to as the *Precise Positioning Service* (PPS). Some of the codes used in the military signals are encrypted. The accuracy in positioning is slightly better in the PPS compared to the SPS. The names *Precise* and *Standard* are left from the time when the accuracy in the civilian part (SPS) was intentionally degraded. This degradation was removed May 1, 2000. The expected error in a position determined by a GPS receiver is about 10 m horizontal and 20 m vertical 95 % of the time.

3.1.1 Satellite Orbits The 24 satellites are divided into six Earth-centered orbital planes with at least four satellites in each plane. The orbits are nearly circular and the six planes are equally spaced along the equator. The inclination angle relative to the equator is about 55°. An illustration of the satellite orbits is shown in Figure 3.1.

To be able to calculate a position, the user must know the coordinates of the satellites. These coordinates can coarsely be calculated from *GPS almanacs* that list the orbit parameters for all satellites. The almanacs are transmitted in the satellite data message. They can also be obtained for off-line use[1]. GPS receivers use the almanacs to determine which satellites that are currently in view. Without this information the

[1] E.g. from `http://www.navcen.uscg.gov/ftp/GPS/archives/yuma/`

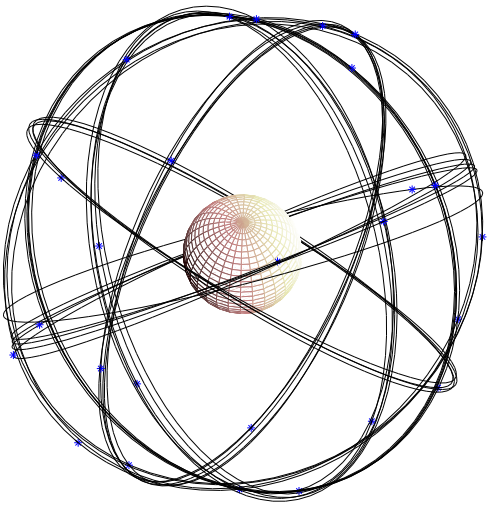


Figure 3.1: GPS satellite movements during 24 hours GPS week 168, November 2002. At this time there were 29 satellites in the constellation. The orbital period for a satellite is about 11 hours and 58 minutes.

receiver can not know which satellites it should try to track. For a finer calculation of the satellite positions the *ephemeris data*, transmitted by the satellites in the data message are used. The ephemeris data contain more accurate parameters which make it possible to determine the satellite position within a few meters.

3.1.2 GPS Signals The positioning with GPS is based on radio signals transmitted from the satellites. The signals transmitted are a mixture of codes and data messages that are needed to determine a distance between a satellite and the receiver and to separate signals from different satellites. The signals can mathematically be written as

$$\begin{aligned} L_1(t) &= A_{p1} p_s(t)\, d(t) \cos(2\pi f_1 t) + A_{c1} c_s(t) d(t) \sin(2\pi f_1 t) \\ L_2(t) &= A_{p2} p_s(t)\, d(t) \cos(2\pi f_2 t) \end{aligned} \tag{3.1}$$

where the two different carrier frequencies are

$$\begin{aligned} f_1 &= 1575.42 \text{ MHz} \\ f_2 &= 1227.60 \text{ MHz} \end{aligned}$$

for the two frequency bands, L1 and L2 in use. A_{p1}, A_{c1} and A_{p2} are signal amplitudes and $p_s(t)$ and $c_s(t)$ are binary spreading codes for satellite s. These codes are known as *Precision* or P-code (p_s) and *Coarse/Acquisition* or C/A-code (c_s). The spreading codes are unique for each satellite. Both the P- and the C/A-code are public, but the P-code in turn is encrypted with a code with regulated access that is only known by authorized users. The last code, $d(t)$ is the 50 Hz data message broadcasted to the receivers. This message contains information about the satellite condition, e.g. clock offsets and orbit parameters. Both codes (P and C/A) and the data message consist of ones ($+1$) and negative ones (-1). From a communication point of view, this is binary phase shift keying (BPSK).

The correlation properties for the codes show similarities with the ones for discrete random signals. That is the reason why these codes are often called *Pseudo Random Noise* (PRN). However, the codes are actually deterministic and can be regenerated exactly. More about the codes used in GPS and how they are generated can be found in [8].

The PRN coded satellite signals are used for two purposes, to separate different signals and to determine the propagation delay. All satellites transmit with the same carrier frequencies f_1 and f_2. Hence, the receiver must be able to separate the messages from the satellites. This is the first thing the C/A- and P-code are used for. The codes are unique for each satellite and have special correlation properties, the cross correlation is almost zero for all time shifts for different codes. This property makes it possible for a receiver to discriminate signals from different satellites. The auto-correlation is also special for the codes. It is 1 for a zero time shift and close to zero for all other time shifts. This property is used in the receiver to determine the propagation time by correlating the received signal with an internally generated copy of the transmitted signal. The time shift for the highest correlation is a measure of the propagation time.

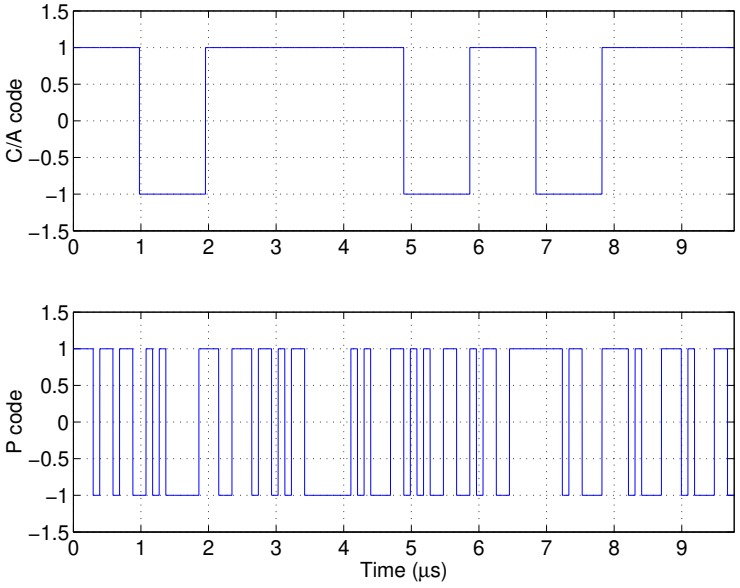
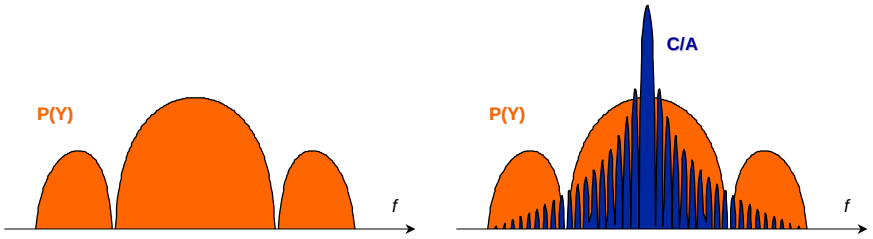


Figure 3.2: Example of codes in GPS. The chip rate is $1.023 \cdot 10^6$ chips/s for the C/A-code and $10.23 \cdot 10^6$ chips/s for the P-code

The PRN C/A- and P-code have different chip rates, $10.23 \cdot 10^6$ chips/s for the P-codes and $1.023 \cdot 10^6$ chips/s for the C/A-codes, see Figure 3.2. The term "chip" is often used for the PRN codes instead of "bit" to emphasize that the codes do not contain any information. The C/A-codes are 1023 samples or 1 ms long while the used part of the P-code is one week long. As mentioned before, each satellite has its own unique C/A- and P-code.

The higher chip rate of the P-code gives a better resolution in the correlation process (since 1 chip corresponds to a shorter time) and therefor a better positioning capability than the C/A-code. Unfortunately, civilian users can, in general, not use this more accurate method since the P-code is encrypted by another unknown code, controlled by the USA military. However, there are methods to circumvent this problem. One method is the so-called Z-tracking, described in [9], where the unknown code is estimated and subsequently removed.

When the data message with a bit rate of 50 Hz is modulated with the more high frequent C/A- and P-codes the energy of the data signal is spread to a wider frequency band, see Figure 3.3. This is the reason why the codes are also referred to as *Spread spectrum codes*. The technology where several users share the same frequency band is generally called *Code Division Multiple Access* (CDMA). The spectras of the GPS signals are depicted in Figure 3.3.



(a) L2, carrier frequency 1227.60 MHz (b) L1, carrier frequency 1575.42 MHz

Figure 3.3: Power spectrum for GPS signals on the L2 and L1 frequency band.

3.2 Pseudorange determination

The signals transmitted by the satellites are used to measure the Time of Arrival (TOA) at the receiver. Since the propagation speed (the speed of light) is known, the distance can be calculated from the delay.

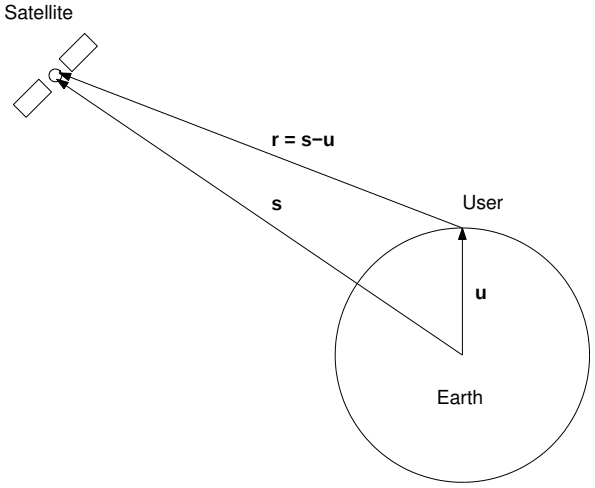


Figure 3.4: Definition of the user-to-satellite vector **r**.

The position of the satellite, **s**, is known from satellite ephemeris data broadcasts (ECEF coordinates). If a user is located at position **u**, the user-to-satellite vector **r** can then be written as (see Figure 3.4)

$$\mathbf{r} = \mathbf{s} - \mathbf{u}. \tag{3.2}$$

Let r represent the distance between a satellite and the user,

$$r = \|\mathbf{r}\| = \|\mathbf{s} - \mathbf{u}\|. \tag{3.3}$$

The distance r, is determined by measuring the propagation time of a signal from the satellite to the receiver. If the true time is T_s when the signal is sent from the satellite and the true time is T_r when the signal reaches the receiver, the geometric distance r can be written

$$r = c\,(T_r - T_s), \tag{3.4}$$

where c is the speed of light.

A big issue when measuring these distances is the clock synchronization. Because of clock offsets, both in the satellite and in the receiver clock, the measured range will be biased and therefore denoted as a *pseudorange*. It contains the true (geometrical) range r, a range offset due to the receiver clock offset and a range offset due to the satellite clock offset. If δt denotes the time the user clock is running fast, and δt^s the corresponding time for the satellite, the measured pseudorange ρ can be written

$$\begin{aligned} \rho &= c\left[(T_r + \delta t) - (T_s + \delta t_s)\right] + \epsilon \\ &= c\left(T_r - T_s\right) + c\left(\delta t - \delta t_s\right) + \epsilon \\ &= r + c\left(\delta t - \delta t_s\right) + \epsilon, \end{aligned} \tag{3.5}$$

where ϵ includes errors, e.g. noise or atmospherical disturbances, affecting the measurement.

The clocks in the satellites are atomic clocks (rubidium or cesium) that are very stable. Nevertheless, these clocks can have a constant deviation δt_s of several milliseconds from GPS system time. To circumvent this problem, the ground control stations are monitoring the satellite clock deviations and uploading correction parameters. These parameters are then transmitted to the GPS receivers via the satellite broadcast data, and the receivers can use this information to compensate for satellite clock errors. In practice the satellite clock offsets after correction are small compared to other errors.

The clocks in the receivers are not as accurate as the satellite clocks. In addition, there is no way of monitoring the clock errors of all GPS receivers in the world. Therefore the receiver clock error δt is the by far most dominating clock error.

3.2.1 Code and Phase Measurement The pseudoranges are determined by correlating the received signal with a receiver generated copy. The properties of the PRN code determine the precision obtained by this correlation. As a complement, there is another, more sophisticated way to determine the pseudorange. This is done by measuring the phase of the received signal before demodulation. It gives a very accurate measure with an uncertainty which is a fraction of a wavelength. The problem is that it is only possible to measure the phase φ of the last fraction of a wavelength, and not the real distance. There will be an unknown number of whole wavelengths N passed, that can not be measured directly, see Figure 3.5. Hence, the measured phase pseudorange can be written as

$$\lambda\,\varphi = r - \lambda\,N. \tag{3.6}$$

The unknown number of wavelengths N is called the *Integer ambiguity*. There are several ways of solving the integer ambiguity, e.g. [10].

When the integer ambiguity is determined, the receiver can track the change in phase due to movements of the receiver and the satellites. Therefore the integer ambiguity must only be determined once as long as the receiver can track the signal.

The integer ambiguity has to be determined for each satellite tracked. The integer is a number in the order $N \approx 10^8$. Without determining the integer ambiguity, the phase measurement is rather worthless. Hence, compared to the code measurement, the phase measure is more accurate but less robust. Typical values for the standard deviation of the noise that affects the measurements are listed in Table 3.1.

It should be noted that also the C/A code measurement contains an ambiguity. This is because the code repeats every millisecond (about 300 km), and the distance to the satellites is much longer. However, this ambiguity is easily resolved by setting a coarse initial position. A position within a few hundred kilometers from the correct one is sufficient [9].

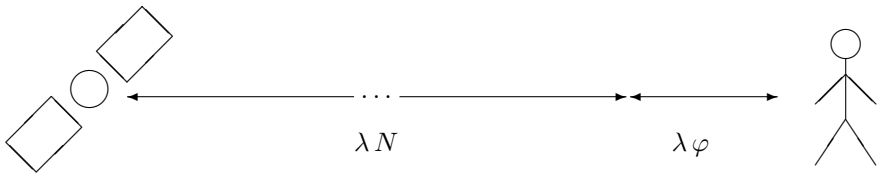


Figure 3.5: In the phase measurement, the phase φ of the received signal is determined. But there is an unknown number of whole wavelengths N passed, that can not be measured directly. Hence, the phase measurement alone does not give any information about the total distance between the satellite and the user.

Measurement	Noise σ	
C/A code	$0.10 - 3$	m
P code	$0.1 - 0.3$	m
Phase	$2 \cdot 10^{-4} - 5 \cdot 10^{-3}$	m

Table 3.1: Typical values of noise standard deviations affecting the pseudorange (from [9]). Note that the more accurate phase measurement can only be used if the integer ambiguity is resolved.

3.2.2 Doppler Measurement Movements of the receiver will give a Doppler shift of the received signals from the satellites. This Doppler shift can be used as a measurement of the receiver velocity in the direction of the satellite. The movement of the satellite, of course, also gives rise to a doppler shift, but the satellite movements are highly predictable, and can be compensated for. Note that it is only the velocity in the direction of the line from the receiver to the satellite, that contributes to the Doppler shift.

The measurements of velocities, or as it is denoted in literature, *delta pseudoranges*, can be used in the same way as the previously described pseudoranges. When it is possible to determine a position in three dimensions from a set of pseudoranges, a velocity in three dimensions can be determined from a set of delta pseudoranges.

3.3 Position Determination

There are various methods of how to calculate the position once the pseudoranges are determined. It can be noted that since the receiver clock offset is unknown and the three-dimensional position contains three unknown parameters, there are a total of four unknown parameters. To solve a system with four unknowns, at least four equations are needed, which in the GPS case corresponds to four pseudoranges. There exist closed form solutions for the position if four pseudoranges are available. However, a better approach, in most cases, is to measure as many pseudoranges as possible. The system then becomes over-determined, but can be solved in a least squares sense.

Let the user's position be $\mathbf{u} = (x, y, z)$ and let $\mathbf{s}_j = (x_j, y_j, z_j)$ denote the position of the j^{th} satellite, both in ECEF coordinates. The satellite position is known from the broadcasted data message. The system with m equations can then be written as

$$\rho_j = \sqrt{(x_j - x)^2 + (y_j - y)^2 + (z_j - z)^2} + c\,\delta t + \epsilon, \quad j = 1, 2, \ldots m. \tag{3.7}$$

Note that these equations are nonlinear.

A geometrical interpretation of the system in (3.7) is illustrative. Each measured pseudorange sets a constraint for how far from the satellite the user is. No information about the direction is given so the constraint forms a sphere, with the satellite in the middle and where the user is somewhere on the surface.

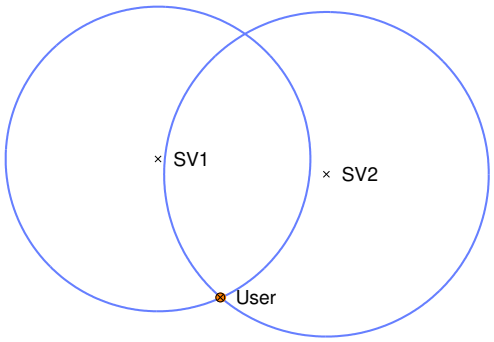
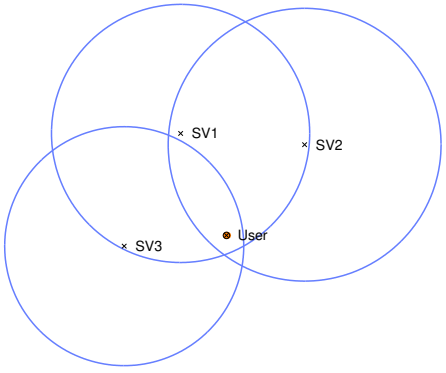
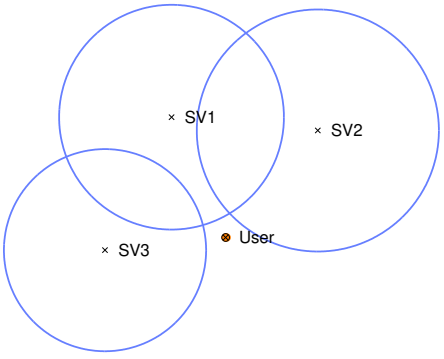


Figure 3.6: Determining a position from a set of pseudoranges in a two-dimensional case.

For a similar, but easier-to-visualize two-dimensional case, each pseudorange forms a circle where the user can be positioned. In an ideal case, with no clock offsets, it should be possible to determine a position by using only two pseudoranges/circles, if a coarse user position is known (see Figure 3.6). If the receiver clock has an unknown offset it is not possible to determine the user position from only two pseudoranges (see Figure 3.7). A third pseudorange giving an extra constraint can solve this problem since the receiver clock offset can be solved for as an unknown parameter. Note that this example is only valid for the two-dimensional case, with two unknown coordinates and one unknown clock offset. To solve for the four unknown in (3.7) at least four pseudoranges must be measured.



(a) A positive receiver clock offset gives over-estimated pseudoranges.



(b) A negative receiver clock offset gives under-estimated pseudoranges.

Figure 3.7: Determining a position from a set of pseudoranges in a two-dimensional case with a non-zero receiver clock offset. A least square solution without taking the clock offset into account would not give the correct position.

The same principle is valid for the delta pseudorange, described in Section 3.2.2. Ideally, if three delta pseudoranges are measured, the velocity in three dimensions can be calculated. The same problem with the receiver clock appears also in the Doppler measurement, so in reality four delta pseudoranges has to be measured to determine a three-dimensional velocity. The last measure is used to determine the drift of the receiver clock, i.e. the frequency offset. This is equivalent with determining the receiver clock offset $c\,\delta t$ in the positioning determination case.

3.4 Error Sources

The measured pseudoranges are affected by several errors. Before May 2000, the largest error was *Selective Availability* (SA). SA is an intentional clock error introduced in the satellites by the US Department of Defence to avoid unauthorized people to get, in their opinion, a too accurate position. The SA signal was removed in May 2000 but there is no guarantee from the USA that it will not be introduced again.

In Table 3.2 typical errors that affect a pseudorange measurement are presented. With the assumption that the errors are uncorrelated, they can be summed in a square root sense, to form the *User Equivalent Range Error* (UERE). The UERE is used as an indication of the reliability of one measurement.

Type of error		Typical size (1-σ)
Ephemeris perturbations	δr_{orb}	4.2 m
Satellite clock error	$c\,\delta t_s$	1.0 m
Selective Availability	δr_{SA}	0 m
Ionospheric delays	δr_{ion}	5 m
Tropospheric delays	δr_{trop}	1.5 m
Multipath errors	δr_{MP}	2.5 m
Receiver thermal noise	v	1.5 m
User Equivalent Range Error	σ_{UERE}	7.4 m

Table 3.2: Typical errors affecting the C/A-code pseudorange measurements [3]. If SA is used δr_{SA} can be set to 32.3 m (1-σ). All errors are assumed to be uncorrelated and the error sources are summed in a square root sense to form the *User Equivalent Range Error* σ_{UERE}.

With the notation from Table 3.2 the measured pseudorange from equation (3.5), can be written as

$$\rho = r + c\,(\delta t - \delta t_s) + \delta r_{ion} + \delta r_{trop} + \delta r_{SA} + \delta r_{MP} + \delta r_{orb} + v \tag{3.8}$$

The term $c\,\delta t$ is the receiver clock offset compared to GPS time. The reason why this error is not taken into account in the σ_{UERE} is that it is usually calculated directly from the measured pseudoranges, as described in Section 3.3.

3.4.1 Atmospherical Effects When the GPS signals propagate towards Earth, they are affected by the atmosphere. The largest effect is due to the *ionosphere* and the *troposphere*.

The propagation speed of a wave through a medium depends on the refractive index n of the medium. For an electromagnetic wave the propagation speed v_p is

$$v_p = \frac{c}{n}, \tag{3.9}$$

where c is the speed of light.

When talking about the propagation speed of an electromagnetic wave we have to distinguish between *phase speed*, which is the propagation speed of the phase for a single frequency wave, and the *group speed*, which is the propagating speed for the energy from a group of waves with slightly different frequencies. Equation (3.9) is valid for both the phase speed and the group speed, but the refractive index n differs.

A medium is said to be dispersive if its refractive index is a function of the frequency of the propagating wave. In the atmosphere, there is a layer between about 70 km and 1000 km, the ionosphere, that is a dispersive medium. If higher order terms are

neglected, the refractive index for the phase speed, n_{ph}, and the group speed, n_{gr}, can according to [3] and [9], be approximated with

$$n_{ph} = 1 + \frac{c_2}{f^2} \tag{3.10}$$

$$n_{gr} = 1 - \frac{c_2}{f^2}, \tag{3.11}$$

where f is the frequency of the wave and the coefficient c_2 depends on the concentration of free electrons N_e in the ionosphere along the propagation path:

$$c_2 = -40.3\, N_e. \tag{3.12}$$

Here N_e is the number of free electrons per m^3. Since the electron concentration N_e is always positive, which implies that $n_{gr} > 1 > n_{ph}$ and further, the group speed, v_{gr}, will always be smaller than the phase speed, v_{ph}: $v_{gr} < c < v_{ph}$. Applying this on the GPS signals says that the phase pseudorange is measured too short and code pseudorange is measured too long, compared to the true geometric range. The concentration of free electrons depends on the time of day. It is largest a few hours after noon and smallest at night.

In the GPS data message transmitted by the satellites, there are parameters for modelling the ionospheric delay. By using this model, the *Klobuchar ionospheric correction model*, approximately 50% of the error can be removed. The model is described in [7].

The neutral troposphere is a non-dispersive medium. Hence, the tropospheric delay does not depend on the frequency.

3.4.2 Dilution of Precision The accuracy of the position and clock offset solution is determined by the uncertainties in the pseudorange measurements, but also by the satellite constellation. The relationship between the pseudorange errors and the computed position and time bias errors are described by the *Dilution of Precision* (DOP), which is a measure of how favorable the satellite constellation is. The DOP is often between, say, 2 and 8, where a low value indicates a good constellation. Examples of constellations that lead to a large DOP are if all satellites are at the same elevation angle above the horizon or located at a straight line from the user's point of view.

An error in the pseudorange measurements $\boldsymbol{\rho}$ gives an error contribution to the position and the receiver clock offset. If the error in the pseudorange $\delta\boldsymbol{\rho}$ can be considered small, the relation between a small error in position $\delta x, \delta y, \delta z$, and in the receiver clock offset $\delta(c\,\delta t)$ and a corresponding small error in the pseudorange $\delta\boldsymbol{\rho}$ can be linearized and written as

$$\delta\boldsymbol{\rho} = \mathbf{H}\,\delta\mathbf{z} \tag{3.13}$$

where $\delta\mathbf{z} = (\,\delta x,\ \delta y,\ \delta z,\ \delta(c\,\delta t))^T$ is a (4×1) vector with the errors in the unknown variables, and $\mathbf{H}$ is the matrix of partial derivatives of the non-linear measurement equation (3.7), with respect to the unknown states:

$$\mathbf{H} = \frac{\partial\boldsymbol{\rho}}{\partial\mathbf{z}} = \begin{bmatrix} -e_{x,1} & -e_{y,1} & -e_{z,1} & 1 \\ -e_{x,2} & -e_{y,2} & -e_{z,3} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ -e_{x,m} & -e_{y,m} & -e_{z,m} & 1 \end{bmatrix} \in \mathbb{R}^{m\times 4} \tag{3.14}$$

where $\boldsymbol{\rho} = (\,\rho_1, \cdots, \rho_m)^T$ are the measured pseudoranges to the m available satellites and

$$(e_{x,j},\ e_{y,j},\ e_{z,j}) = \frac{\mathbf{s}_j - \mathbf{u}}{\|\mathbf{s}_j - \mathbf{u}\|},\quad j = 1, \cdots, m, \tag{3.15}$$

is a unit vector $\mathbf{e}_j$ pointing from the user position $\mathbf{u}$ towards the position $\mathbf{s}_j$ of the j^{th} satellite. The least squares solution to (3.13) for $\delta\mathbf{z}$ can now be written

$$\delta\mathbf{z} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\,\delta\boldsymbol{\rho} \tag{3.16}$$

Pseudorange errors $\delta\boldsymbol{\rho}$ are assumed to be independent and Gaussian distributed with variance σ_ρ^2 and $E[\delta\boldsymbol{\rho}] = \mathbf{0}$. Since matrix multiplication is a linear operation the position and time error $\delta\mathbf{z}$ will also be Gaussian distributed with $E[\delta\mathbf{z}] = \mathbf{0}$. The covariance of $\delta\mathbf{z}$ is

$$\begin{aligned}\mathrm{cov}(\delta\mathbf{z}) &= E[\,\delta\mathbf{z}\delta\mathbf{z}^T] \\ &= E[\,(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\delta\boldsymbol{\rho}\delta\boldsymbol{\rho}^T((\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T)^T] \\ &= (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathrm{cov}(\delta\boldsymbol{\rho})\,\mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\end{aligned} \tag{3.17}$$

Since $\boldsymbol{\rho}$ is assumed to be independent with the same variance σ_ρ^2, the center term in (3.17) can be written as

$$\mathbf{H}^T\mathrm{cov}(\delta\boldsymbol{\rho})\mathbf{H} = \sigma_\rho^2\mathbf{H}^T\mathbf{H}, \tag{3.18}$$

which implies that (3.17) can be simplified to

$$\mathrm{cov}(\delta\mathbf{z}) = (\mathbf{H}^T\mathbf{H})^{-1}\sigma_\rho^2. \tag{3.19}$$

The matrix $(\mathbf{H}^T\mathbf{H})^{-1}$ is always of size (4×4):

$$\left(\mathbf{H}^T\mathbf{H}\right)^{-1} = \begin{bmatrix} D_{11} & D_{12} & D_{13} & D_{14} \\ D_{21} & D_{22} & D_{23} & D_{24} \\ D_{31} & D_{32} & D_{33} & D_{34} \\ D_{41} & D_{42} & D_{43} & D_{44} \end{bmatrix} \in \mathbb{R}^{4\times 4} \tag{3.20}$$

The diagonal elements indicate how the measurement errors in the pseudoranges are projected to an error in the calculated position and clock offset.

The following DOP parameters can be extracted from the diagonal of (3.20):

$$PDOP = \sqrt{D_{11} + D_{22} + D_{33}} \tag{3.21}$$

$$GDOP = \sqrt{D_{11} + D_{22} + D_{33} + D_{44}} \tag{3.22}$$

$$TDOP = \sqrt{D_{44}} \tag{3.23}$$

Here $PDOP$ is the *position* dilution of precision, $GDOP$ is the *geometric* dilution of precision and $TDOP$ is the *time* dilution of precision. The three parameters in (3.21)-(3.23) are valid in any coordinate system. Two more DOP parameters can be defined which are only valid in a local system. If the local system describes a coordinate in the order [East, North, Down] (a so-called ENU system), the two parameters are:

$$HDOP = \sqrt{D_{11} + D_{22}} \tag{3.24}$$

$$VDOP = \sqrt{D_{33}}, \tag{3.25}$$

where $HDOP$ is the *horizontal* dilution of precision and $VDOP$ is the *vertical* dilution of precision. In this case it is sufficient that the matrix $\mathbf{H}$ in (3.14) gives the direction to the satellites in ENU coordinates.

The DOP is a relation between the standard deviation $\sigma_{\boldsymbol{\rho}}$ for the error in the pseudorange measurements and the standard deviation for the error $\sigma_{\mathbf{z}}$ in the determined position and clock offset solution. Which components that the errors are projected on is determined by which of the DOP parameters defined in (3.21)–(3.25) that is used. E.g. $GDOP$ gives the standard deviation for the position error and $TDOP$ the same but for the clock offset. The relation is

$$\sigma_X = XDOP \cdot \sigma_{\boldsymbol{\rho}}, \tag{3.26}$$

where $XDOP$ is any of the previously defined DOP parameters and σ_X is the standard deviation for the errors in the components of $\mathbf{z}$ determined by the chosen DOP parameter $XDOP$. In practice the standard deviation for the pseudorange error is unknown. Often $\sigma_\rho = \sigma_{UERE}$ is used, where the *User Equivalent Range Error* (UERE), σ_{UERE}, is the resulting standard deviation for errors in the pseudorange, if all errors are approximated with uncorrelated Gaussian distributions. Typical errors are listed in Table 3.2. Note that the DOP is determined by the satellites and user positions only.

If less than four satellites are visible, the matrix $\mathbf{H}$ will have less than four rows and hence $\mathbf{H}^T\mathbf{H}$ will not have full rank. Therefore the inverse $(\mathbf{H}^T\mathbf{H})^{-1}$ taken in (3.20) will not exist, which indicates that no position can be calculated from less than four satellites in a single measurement. However, the information from the visible satellites, even if they are less than four, can be used if the GPS and the INS are integrated in a tightly coupled filter. This is one of the main advantages with the tight INS/GPS integration.

Commercial GPS receivers can often present some type of DOP, where $HDOP$ and $VDOP$ are the most commonly used.

3.5 Data Combinations

The pseudorange measurements in the GPS receiver can be combined or manipulated to forms where some of the errors described in 3.4 are reduced. Since the SA was turned off, the largest errors are disturbances from the atmosphere. Most of the data combination methods therefore try to minimize these errors.

3.5.1 Differential Measurement One of the most common types of data combinations is the differential measurements. By differentiating between measurements, errors that are common in the two measurements can be reduced or even eliminated. For example the atmospheric disturbance is very similar in two receivers if they are close to each other.

Single Difference The most common differentiating method is to differentiate between two receivers. This is normally called *single difference GPS* or just *Differential GPS* (DGPS). DGPS achieves enhanced accuracy since the two receivers experience common errors that can be removed if measurements from both receivers are available.

Usually one of the two receivers is at rest at a position with known coordinates, while the other one is roving. Since the receiver at rest, often referred to as a base station, knows its coordinates, it knows what pseudoranges it should expect to measure. When the real erroneous pseudorange is measured a pseudorange correction can be calculated. The correction can be broadcasted (e.g. via a radio link) to all roving GPS receivers. The roving receiver can measure the pseudoranges and correct them by using the received correction term calculated by the base station.

Using differential GPS techniques can increase the accuracy considerably. This is because the GPS error sources are very similar over a short distance and are therefore virtually eliminated. For baselines less than 20 km GPS errors such as satellite clock errors, orbital errors, Selective Availability, ionospheric and tropospheric delays can be considered constant and if they are removed accuracies in sub-decimeter level can be obtained.

One error that can not be eliminated with differential corrections is the thermal noise in the receiver, v. Instead, its variance increases with a factor two. This is under the assumption that the noise in the base station and the receiver is uncorrelated, which

is reasonable.

When the dithering of the satellite clocks with the SA signal was used, the technology with DGPS gave a more significant precision improvement, since this error was effectively removed by the differential correction.

Double Difference The method of double difference uses single differenced data and in addition takes the difference between measurements from different satellites. Since single differenced data are used, atmospheric effects, orbit errors, SA and satellite clock errors are already eliminated. In addition, errors common to the two satellites are eliminated. One common error is the receiver clock offset. The drawback with taking the difference is, as in the single difference case, that the uncorrelated noise is increased. Since the receiver clock offset is easily estimated, double difference will not be used in this thesis.

Triple Difference A triple difference uses two double differenced measurements from two different times. Again, common errors, which now are the integer ambiguities are eliminated.

3.5.2 Dual Frequency As described in Section 3.4.1, the ionosphere has the main atmospheric effect on the measured pseudoranges. Due to the dispersive nature of the ionosphere, where signals are affected depending on the frequency, it is possible to eliminate its effect if the pseudoranges are measured with two different frequencies [4] [9].

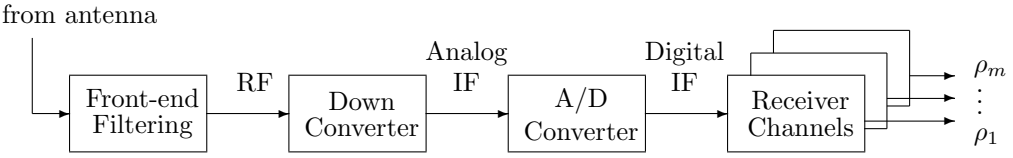


Figure 3.8: A high-level block diagram of a typical digital GPS receiver.

3.6 Receiver Architecture

To analyze how interference and noise affect the measurements, a more detailed description of the GPS receiver design is needed. Today almost all GPS receivers are digital, which means that the signal processing is performed with sampled digital signals. A high-level block diagram describing the construction of a digital GPS receiver is shown in Figure 3.8.

The radio frequency (RF) signals transmitted by the satellites are received by an antenna. First, the received signal is filtered in a front-end filter to remove out of band noise and interference. The analog signal is then down-converted to an intermediate frequency (IF) signal. In the down-conversion the PRN codes and the Doppler information in the signal are preserved, only the frequency of carrier wave is decreased [3]. The frequency of the IF signal is chosen to be sufficient for signal processing, usually a few megahertz. Not shown in the block diagram is the anti-aliasing filter that cuts the upper side-band of the down-converted signal. In the next step the IF signal is sampled to a digital signal. The digital signal is then processed in the receiver channels to calculate pseudoranges $\boldsymbol{\rho}$ to the satellites in view. In receivers bought today,

the processing is often done in parallel as indicated in the block diagram. Receivers with 12 channels in parallel are common. A 12 channel receiver can simultaneously track up to 12 satellites.

3.6.1 Tracking Loops Measuring of pseudoranges are done by correlating the received PRN code with an internally generated copy of the code. The offset for the highest correlation is a measure of the propagation time and, hence, also the distance. The pseudoranges will change over time, due to movement of the receiver and the satellites. Therefore the receiver must continuously adjust the offset of the internally generated code to get a maximum correlation with the receiver signal. This is performed in the *tracking loops*. There are two tracking loops in a receiver, a *carrier tracking loop* and a *code tracking loop*. A block diagram of the carrier and the code tracking loops is depicted in Figure 3.9.

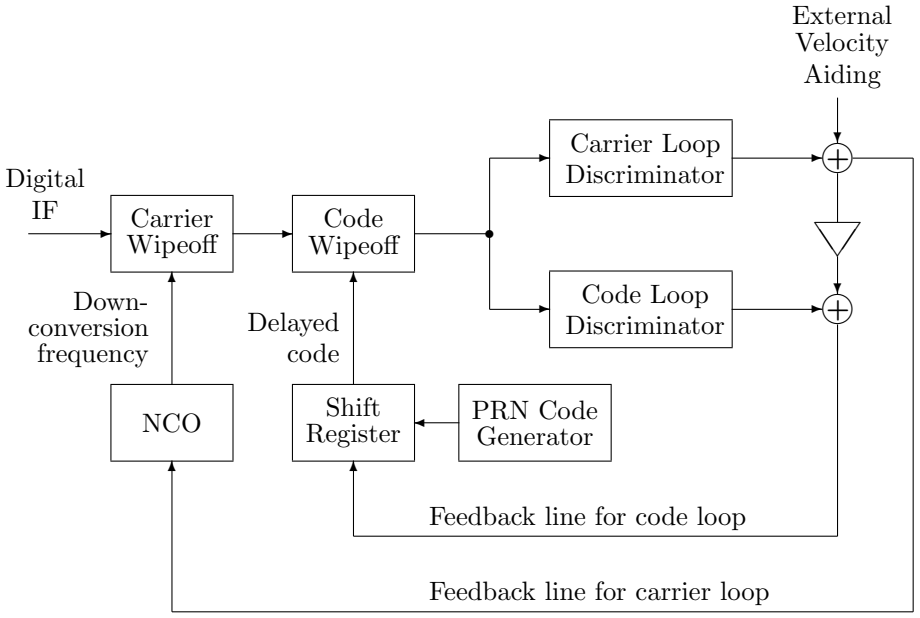


Figure 3.9: Receiver channel containing the code and carrier tracking loops. The carrier tracking loop is aided by an external velocity measurement. The code tracking loop is aided by Doppler data from the carrier tracking loop.

The objective of the carrier tracking loop is to adjust the frequency of the signal used for down-converting. This is necessary because the received signal can be Doppler shifted due to movements of the receiver. For slowly moving vehicles the largest Doppler shift is caused by the movements of the satellites. If the frequency of the received signal is changed, the *carrier loop discriminator* will sense that the frequency is wrong and try to adjust the down-conversion frequency.

During periods with large dynamic stress, the frequency might need large adjustments. There are two strategies to make this possible. Either to make the bandwidth of the carrier tracking loop large, so it can feed large adjustments back to the *Numerically Controlled Oscillator* (NCO), or to aid the carrier tracking loop with an external velocity measure. The latter can be done with an INS. If the carrier tracking loop is locked on the correct down-conversion frequency, the carrier frequency of the Digital IF is removed by multiplying the two signals in the *Carrier Wipeoff*. What is left is a complex baseband signal, representing the data signal, spread by the PRN code.

A carrier tracking loop can be implemented either as a *phase locked loop* (PLL) or a *frequency locked loop* (FLL). An FLL is less sensitive to jamming and dynamic stress compared to a PLL. A PLL, on the other hand, gives more accurate Doppler measurements [3].

The objective of the code tracking loop is to synchronize the received coded signal with an internally generated signal. Thanks to the correlation properties of the PRN codes, the *code loop discriminator* can determine if the output from the *PRN Code Generator* is early or late compared to the received signal and adjusts the delay in a shift register. If the code loop is locked, i.e the correct delay shift is found, multiplication with the known PRN code in the *code wipeoff* despreads the baseband signal, and only the data message and noise remains. When the code loop is locked the delay shift is a measure of the pseudorange. The code tracking loop is often aided by Doppler measurements from the carrier tracking loop.

The despreading process that is performed in the *code wipeoff*, decreases the bandwidth of the signal in the same extent as the amplitude is increased. However, the total power remains the same. The amount that the amplitude is increases is called the *processing gain*. The processing gain, theoretically, is 43 dB for the C/A code[2] and 53 dB for the P code. The two tracking loops imply that the tracking of a signal is a two-dimensional search process. Only if both the down-conversion frequency and the code offset are correct, a pseudorange can be determined.

An important property of the tracking loops, both the carrier and the code loop, is that if their bandwidth is decreased it gives an increased SNR in the loops. A higher SNR gives better tracking accuracy, so the obvious action is to decrease the loop bandwidths. However, there is a drawback. With more narrow tracking loop bandwidths the receivers ability to track Doppler shifts, which arise from the dynamics of the receiver is reduced. A way to narrow the loop bandwidth without limiting the tracking performance is to aid the carrier loop with external velocity measurements [12]. The information of the navigation platform velocity is used to make a coarse track of the carrier phase, then the carrier tracking loop performs the fine track around that predicted value.

Tracking Thresholds The tracking loops need a certain signal strength compared to interfering signals and noise to be able to track the satellite signal. If the SNR is below this value, the *tracking threshold*, the receiver will not be able to present any measurements. It is also necessary that the tracking loops have acquired the signals, which means that both the correct carrier Doppler shift and code offset are found. The *acquisition threshold* is usually a few decibels higher than the tracking threshold. An example of this is shown in Figure 3.10.

The first part in a GPS receiver to lose track due to low SNR is normally the carrier tracking loop [3]. It is therefore interesting to get an idea of how low the SNR can be, without losing lock in the carrier loop. The calculation of tracking thresholds is treated in [3].

The PLL is non-linear and will lose lock for errors larger than 45°. Considering this as a 3-σ limit gives a rule-of-thumb for the 1-σ error:

$$\sigma_{PLL} \leq 15^\circ. \tag{3.27}$$

In [3] σ_{PLL} is approximated with

$$\sigma_{PLL} = \sqrt{\sigma_t^2 + \sigma_v^2 + \sigma_A^2} + \frac{\theta_e}{3}, \tag{3.28}$$

[2]Some authors [11] claim that the process gain for the C/A code in practice is not more than about 24 dB due to the short sequence and therefore bad frequency spreading.

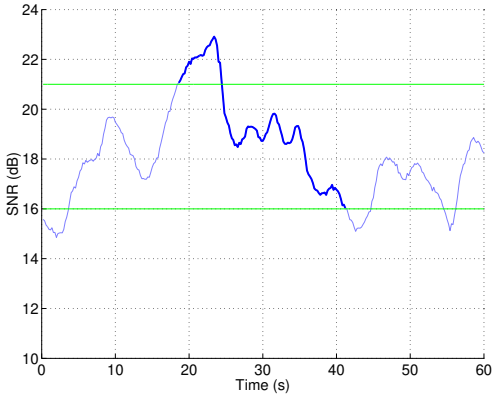


Figure 3.10: The GPS receiver needs a certain SNR to be able to use a satellite signal. In this example the tracking threshold is 16 dB and the acquisition threshold is 21 dB. The receiver can use the signal in the highlighted section, between approximately 18 s and 41 s. These thresholds, 16 dB and 21 dB are the ones that are actually used in the simulations described later. The assumptions leading to these values are listed in Table 8.1.

where

σ_t	=	standard deviation of thermal noise in degrees
σ_v	=	standard deviation of vibration-induced oscillator errors in degrees
σ_A	=	Allan variance-induced oscillation error in degrees
θ_e	=	dynamic stress error in PLL tracking loop,

The thermal noise with standard deviation σ_t is often the dominant term [3]. The most important to note with the thermal noise is that it increases with a wider loop bandwidth B_n. This is the reason for trying to narrow the loop bandwidth by e.g. external velocity aiding.

Finding the vibration-induced oscillator error σ_v, is a complex problem. There are closed form expressions [3], but they require that the characteristics of the oscillator are known.

The Allan variance-induced oscillator error σ_A, also depends on the oscillator characteristics, and no value can be calculated for σ_A in general. It is worth noting however, that σ_A depends on $1/B_n$, where B_n is the bandwidth of the PLL and thus also the incoming noise. Decreasing the bandwidth of the PLL gives a lower thermal noise σ_t but, hence also increases the σ_A. For very low PLL bandwidths, σ_A can be the dominating error. With an unaided PLL, and hence a larger PLL bandwidth, σ_A is normally below 1°.

3.7 Interfering a GPS Receiver

The power of the received signals from the satellites are extremely low because of the long distance to the satellites. Because of the spectrum-spreading property of CDMA, the received GPS signals are actually far below the thermal noise floor [3]. In the despreading process the signals are amplified with the processing gain of the code used.

This makes it possible to make a GPS receiver useless with a relatively low jamming power. Therefore the risk for jamming or unintentional interference is a big issue when designing GPS receivers, especially for military use.

There are two main types of intentional interference that a receiver can be exposed to, jamming and spoofing. The objective with jamming is only denial of service, by drowning the GPS signal in noise so that the receiver tracking loops lose lock. Spoofing means that an incorrect GPS signal is transmitted, that looks like a real signal to the user. The objective can be to make the user think he is at another position than he actually is or to change the ephemeris data, e.g. saying that all satellites are out of order. Jamming is relatively easy and inexpensive. Spoofing is more difficult and requires much more sophisticated equipment. In this report only the effects of jamming is considered.

The received interfering power is depending on the distance from the source. If the most simple propagation model, free-space propagation [13], is assumed, the received power density S (power per unit area) decreases as

$$S = \frac{P_t}{4\pi\, r^2} \tag{3.29}$$

where P_t is the totally transmitted power and r is the distance from the source. How much of the power density that can be transformed into power is determined by the size of the antenna, the *effective aperture area*, that collects the energy. If the effective aperture area is A_e the received power is

$$\begin{aligned} P &= A_e \cdot S \\ &= A_e \cdot \frac{1}{4\pi\, r^2} \cdot P_t. \end{aligned} \tag{3.30}$$

According to [13] the following relation between the effective aperture area and the *antenna gain* G holds

$$G = A_e \cdot \frac{4\pi}{\lambda^2}, \tag{3.31}$$

or equivalently

$$A_e = G \cdot \left(\frac{4\pi}{\lambda^2}\right)^{-1}, \tag{3.32}$$

where λ is the wavelength of the received signal. Using this in (3.30) gives

$$\begin{aligned} P &= G \cdot \left(\frac{4\pi}{\lambda^2}\right)^{-1} \cdot \frac{1}{4\pi\, r^2} \cdot P_t \\ &= \frac{P_t \cdot G}{\left(\frac{4\pi\, r}{\lambda}\right)^2}. \end{aligned} \tag{3.33}$$

In radio communication applications, and so also for GPS, most quantities are given in decibel (dB). The relation between decibel and ratio, e.g. for two power levels P_1 and P_2, is

$$\left[\frac{P_1}{P_0}\right]_{\mathrm{dB}} = 10 \log \left(\frac{P_1}{P_0}\right) \tag{3.34}$$

where "log" is the base-10 logarithm.

With quantities in decibels, the relation in (3.33) can be written as

$$P = P_t + G - L, \tag{3.35}$$

where $L = 20 \cdot \log \left(\frac{4\pi\, r}{\lambda}\right)$ is the *path loss* which corresponds to the denominator in (3.33).

3.7.1 Effects of Jamming The signal-to-noise ratio (SNR) is defined to be

$$SNR = \left[\frac{P_S}{P_N}\right]_{\text{dB}} = 10 \log \left(\frac{P_S}{P_N}\right), \tag{3.36}$$

where P_S and P_N are the received powers of the GPS signal and the noise, respectively.

In GPS applications the carrier-to-noise ratio (C/N_0) is often used instead of the signal-to-noise ratio. It is the ratio between the total carrier power and the noise power over 1 Hz bandwidth. Hence, the unit for C/N_0 is Hz or, if expressed in decibel, dBHz.

In a best case scenario, when the receiver is not exposed to any jamming, the C/N_0 is the ratio between the carrier power and the thermal noise. Using (3.35) and subtracting the thermal noise spectral density and general losses gives

$$\frac{C}{N_0} = (P_{GPS} + G_{GPS} - L_{GPS}) - (10 \cdot \log (k\, T_0)) - L_s \tag{3.37}$$

where

P_{GPS}	=	Transmitted GPS signal power (dBW)
G_{GPS}	=	Antenna gain in direction towards satellite (dB)
L_{GPS}	=	Path losses (dB)
k	=	Boltzmann's constant ($1.38 \cdot 10^{-23}$ Ws/K)
T_0	=	Thermal noise reference temperature (K)
L_s	=	Other losses, e.g. polarization or implementation losses (dB)

According to the GPS specification, referred to in [3], the received GPS signal power $P_{GPS} = P_{GPS} - L_{GPS}$ is guaranteed to be at least -159.6 dBW for the C/A-code.

When the receiver is exposed to interfering signals the C/N_0 decreases. This lower level is called the equivalent carrier-to-noise ratio and is in some literature denoted $[C/N_0]$. If the so-called jammer-to-signal ratio J/S in decibel is defined as

$$J/S = J_r - S_r, \tag{3.38}$$

where J_r is the received jamming power in dBW and S_r is the received signal power in dBW, the $[C/N_0]_{eq}$ can, according to [3], be calculated as

$$[C/N_0] = \left(\frac{1}{10^{\frac{C/N_0}{10}}} + \frac{10^{\frac{J/S}{10}}}{Q\, R_c}\right)^{-1}. \tag{3.39}$$

Here R_c is the chip rate for the code used (e.g. $1.023 \cdot 10^6$ chips/s for the C/A-code) and Q is a dimensionless constant that is decide by the type of jammer. For wideband noise jammers $Q = 2$.

Combining the result of the current $[C/N_0]$ with knowledge about the tracking thresholds (described in Section 3.6.1) gives an idea of when the GPS receiver can or can not be used, because of interfering signals.

4. Adaptive Beamforming for GPS

The extremely low power levels from the satellites make the GPS system vulnerable to jamming and interference. The guaranteed signal power is only −160 dBW, so an interfering signal does not have to be strong to drown the GPS signal.

Interfering signals with powers of several tens of decibels higher than the satellite signals should be expected. Therefore it is often not good enough to form a narrow beam towards the satellite. A better result can be achieved if the null gains in the reception pattern are formed to get a sufficient suppression in the directions of interfering signals.

The interference can be either unintentional from other radio frequency equipment or arise from hostile jamming. The use of antenna arrays and adaptive beamforming is a promising method to overcome this problem. Adaptive beamforming has the ability to change the antenna pattern to take advantage of the current signal environment, and discriminate between signals separated in space. Signals not of interest (SNOI), i.e. jamming or interfering signals, are suppressed, while a main beam is formed towards the signal of interest (SOI) to amplify the signal. An example of this is shown in Figure 4.1.

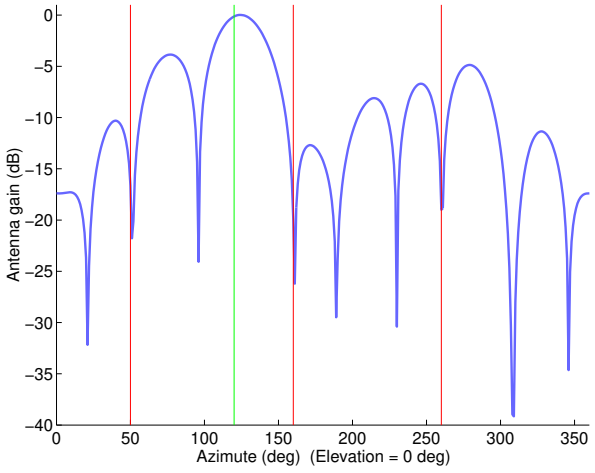


Figure 4.1: A sample adaptive antenna reception pattern. The pattern shows the antenna gain in different directions. A higher gain is set in the direction from where the interesting signal is arriving (in this case 120°). Interfering signals approaching from 50°, 160° and 260° are suppressed by the beamforming algorithm.

There are non-adaptive antennas that can avoid many of the problems in GPS applications, e.g. low SNR or multipath. Antenna elements with a directional (but fixed) antenna pattern can suppress interference from a known direction, for example choke ring antennas. However, an adaptive beamforming antenna is much more flexible. With the use of certain algorithms, an adaptive antenna can also help to prevent multi-path problems by suppressing the reflected components [14]. With some beam-

forming algorithms some protection against spoofing can also be expected.

Compared to a conventional antenna, an antenna array is more complex, often bulkier and more expensive. These drawbacks are compensated for with a significantly better performance in an interfered environment.

4.1 Antenna Array

An antenna array consists of a number of antenna elements, put together in a known geometry. The geometry is usually symmetric, e.g. Uniform Linear Arrays (ULA), Uniform Rectangular Arrays (URA) or Uniform Circular Arrays (UCA). The beam pattern from a ULA is always rotational symmetric, so the power can only be directed in one dimensions. To be able to control the beam pattern in two dimensions, a URA or a UCA with the elements extended in two dimensions must be used.

The smallest mutual distance between elements is usually equal to or slightly less than half a wavelength $\lambda/2$, of the SOI carrier. A larger distance than $\lambda/2$ leads to a kind of "spatial under-sampling". With a smaller distance, the risk for mutual coupling between elements increases. This demand sets a physical limitation on how small the array can be or how many elements that can be used. E.g. in a GPS application a 4×4 element URA can not be smaller than $3 \cdot \lambda/2 \times 3 \cdot \lambda/2$. For the L1 band this is approximately 0.3 m $\times$ 0.3 m, neglecting the size of the antenna elements. The minimum distance sets a upper limitation on how many elements that can be used on a particular navigation platform.

4.2 Adaptive Beamforming Theory

The performance of adaptive beamforming strongly depends on the number of elements in the array. More elements gives a better performance, but also increases the computational complexity. Generally, an N element array has N degrees of freedom. These degrees can be spent on forming a sharp beam in the spatial domain, or to put, at most, $(N-1)$ nulls in directions of unwanted signals.

The incoming signal to the antenna is an analog RF signal. The beamforming will be performed digitally, so all manipulations will be performed on the corresponding time-discrete complex baseband signal $\mathbf{u}(k)$. The complex baseband representation of a bandpass signal is described in [15].

The carrier frequencies of the GPS signals are $f_1 = 1575.42$ MHz (C/A and P) $f_2 = 1227.60$ MHz (only P). The bandwidth of the signals is 20 MHz for the P-code signal and 2 MHz for the C/A-code signal. This makes the relative bandwidth less than 2 % and less than 0.2 % of the carrier frequencies for the P- and C/A-code signals, respectively.

In beamforming theory, a rule-of-thumb for a signal to be considered narrowband is that it should not cover more than $1-2$ % of the carrier frequency, which is a borderline case for the P coded signal. The C/A coded signal can definitely be considered to be a narrowband signal. The general case with a wideband signal requires that a tapped-delay line (FIR filter) is applied behind each element for selection or suppression of the desired frequency.

Since the C/A-code signal are considered narrowband, delays can be described with phase shifts. It is further assumed that the source is far away, which gives plane wavefronts and no loss of signal strength over the antenna array. A signal incident at the array then gives a signal vector

$$\mathbf{u}(k) = [\, u_1(k) \;\; \ldots \;\; u_N(k)]^T \tag{4.1}$$

on the N elements of the array. These signals from the antenna elements are multiplied with a vector of complex weights $\mathbf{w}$

$$\mathbf{w} = [\, w_1 \;\; \ldots \;\; w_N]^T. \tag{4.2}$$

The weights $\mathbf{w}$ is a vector of complex numbers, so that each multiplication corresponds to an amplitude scaling and a phase shift (see figure 4.2). The weight vector $\mathbf{w}$ can

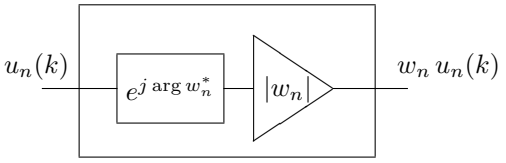


Figure 4.2: An element w_n in the weight vector $\mathbf{w}$, can be seen as a phase shift, $\arg w_n^* = -\arg w_n$, and an amplitude scaling, $|w_n|$.

be changed adaptively depending on the received data. As illustrated in Figure 4.3, the scalar output, $z(k)$, can be written as

$$z(k) = \mathbf{w}^H \mathbf{u}(k), \tag{4.3}$$

where the superscript $(\cdot)^H$ denotes the *Hermitian transpose*[1].

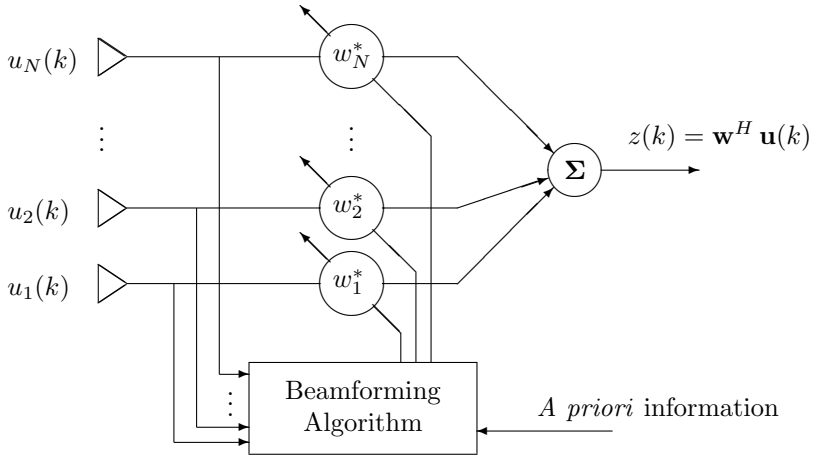


Figure 4.3: A block diagram of an adaptive beamformer for an N-element antenna array.

The direction of a signal arriving from azimuth ϕ and elevation θ can be described with a unit vector $\mathbf{e}$ in Cartesian coordinates, pointing from the array towards the signal source. Azimuth ϕ is defined to be the angle from the y-axis in the plane spanned by the x and y-axes. Elevation is defined as the angle between the x-y plane and the coordinate, with positive direction towards the z-axis. Azimuth and elevation are also illustrated in Figure 4.4. With these definitions the direction unit vector $\mathbf{e}$ in Cartesian coordinates can be written:

$$\mathbf{e}(\phi, \theta) = (e_x, e_y, e_z) = (\sin\phi\cos\theta, \cos\phi\cos\theta, \sin\theta)^T. \tag{4.4}$$

If one of the elements in the array serves as a reference element, say element 1, we can define the difference in propagation distance d_n for a signal between element n and the reference element. The distance d_n can be written as

$$d_n = \mathbf{e}^T \cdot (\mathbf{r}_n - \mathbf{r}_1), \quad n = 1, \ldots, N, \tag{4.5}$$

[1] A complex conjugation combined with transposition, $\mathbf{A}^H = (\mathbf{A}^*)^T$

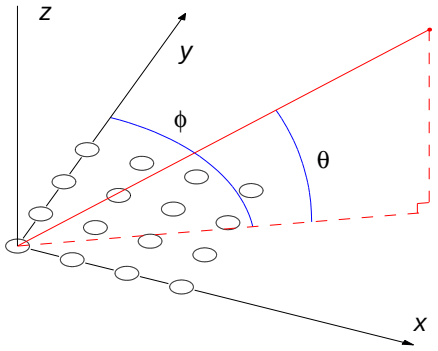


Figure 4.4: Example of a URA element structure in a Cartesian antenna coordinate system. The definitions of the azimuth and elevation angles for a coordinate are shown.

where $\mathbf{r}_n$ and $\mathbf{r}_1$ are the coordinates of the n^{th} element and the reference element, respectively. Hence, d_n is the projection of $(\mathbf{r}_n - \mathbf{r}_1)$ on the direction unit vector $\mathbf{e}$, pointing towards the signal source. The signals in all N elements, compared to the signal in the reference element, can then be written as a *steering vector* [16] [17] with size $(N \times 1)$:

$$\mathbf{s}(\mathbf{d}) = (e^{-j\frac{2\pi}{\lambda}d_1}, e^{-j\frac{2\pi}{\lambda}d_2}, \cdots, e^{-j\frac{2\pi}{\lambda}d_N})^T, \tag{4.6}$$

where $\mathbf{d} = [d_1, \cdots, d_N]^T$ is the vector containing all the propagation difference distances. The steering vector describes the phase shifts in the elements that the requested signal direction gives rise to.

4.3 Comparing Beamforming Algorithms

If the positions of the satellites relative to the antenna are known it would be possible to direct the highest gain of the antenna in these directions. This also requires that the position and orientation of the antenna are known.

The adaptive beamformer is designed to be used with an integrated INS/GPS. Two of the outputs from the integrated system are the position and attitude relative to Earth. These outputs are, thanks to the INS, determined much more accurate than with a stand-alone GPS. From this, the directions to the GPS satellites can be calculated and used as *a priori* information to the beamforming algorithm. With an integrated INS/GPS, estimates of position and attitude are available also during periods with satellite outages.

A wide range of different adaptive beamforming algorithms have been proposed in the literature. All algorithms utilizes some kind of *a priori* information to discriminate between the SOI and jammer signals, or *signals not of interest* (SNOI). The *a priori* information in a GPS application can be:

- The direction of arrival, either for the SOIs or the SNOI.
- Correlation properties of the GPS signal or the spreading codes.

With the *a priori* information above, there are a few algorithms that are possible to use.

4.3.1 Correlation Based Algorithms If a known reference signal is appended to the SOI, correlation methods can be used to find the DOA. The algorithm tries to minimize the difference between the actual received signal and the reference signal.

This require that a known sequence is appended to the signal. Unfortunately, this is not the case with the GPS signal today [14], so a reference signal based method is not possible to use.

An additional problem is that correlation based algorithms are vulnerable to interference by signals similar to the reference signal. A person who wishes to interfere a receiver using a reference signal based beamforming antenna, could simply retransmit a copy of a PRN code. The beamformer could then mistake the interfering signal for the one from the satellite and put a large gain in the wrong direction.

It may also be possible to use the known spreading codes for correlation. However, the problem with false signals that look like the GPS signals remains.

4.3.2 Direction Constrained Null Steering If the DOAs of interfering signals are known, or can be estimated, the antenna pattern can be adapted to form genuine nulls in these directions [2]. As mentioned before, an N element antenna array can form a maximum of $N-1$ nulls in different directions.

The main problem with this algorithm is to estimate the DOAs for the, at most $(N-1)$, strongest interferers. If less than $(N-1)$ interferers exist, a difficulty can be to estimate this number. Using a direction constraint null steering algorithm is a little bit of changing one problem, forming the weight vector, to another possibly harder problem, namely finding the direction of an unknown number of interferers.

4.3.3 Minimum Variance Algorithm Minimum variance algorithms minimize the power of the received signal at the array output z. Smaller gains are set in directions where interferers are present, which gives effective suppression of strong interfering signals. The minimum variance algorithm can actually work without any *a priori* information, but there is a risk that also the SOI is suppressed. To avoid that the SOIs are suppressed it is common to set constraints for the gain in the directions of the SOIs.

The minimum variance condition results in a expression for the weight vector $\mathbf{w}$ as

$$\begin{aligned} \min_{\mathbf{w}} \varepsilon[\,|z(k)|^2] &= \min_{\mathbf{w}} \varepsilon[\,z(k)\,z^H(k)] \\ &= \min_{\mathbf{w}} \varepsilon[\,\mathbf{w}^H\mathbf{u}(k)\,(\mathbf{w}^H\mathbf{u}(k))^H] \\ &= \min_{\mathbf{w}} \mathbf{w}^H \varepsilon[\,\mathbf{u}(k)\mathbf{u}^H(k)]\,\mathbf{w}. \end{aligned} \tag{4.7}$$

where $\varepsilon[\cdot]$ denotes the time average. The factor $\varepsilon[\,\mathbf{u}(k)\mathbf{u}^H(k)]$ is the *array correlation matrix* and is denoted with $\mathbf{R}(k)$:

$$\mathbf{R}(k) \equiv \varepsilon[\,\mathbf{u}(k)\mathbf{u}^H(k)]. \tag{4.8}$$

The array correlation matrix contains the spatial correlation of the incoming signals in different antenna elements.

The gain constraints are set as

$$\mathbf{S}^H\mathbf{w} = \mathbf{g}, \tag{4.9}$$

where $\mathbf{S} = [\,\mathbf{s}_1, \ldots, \mathbf{s}_L]$ is a *steering matrix*, i.e. a matrix containing L steering vectors if L constraints are set. The vector $\mathbf{g}$ defines the fixed gains in the constraint directions. Each of the L vectors in $\mathbf{S}$ forms one constraint. When gain constraints are set for a minimum variance algorithm, it is often called *Linear Constrained Minimum Variance* (LCMV), which is the term that will be used in this thesis.

The solution to the condition (4.7), using the constraint in (4.9), is the optimal weight vector using the LCMV algorithm

$$\mathbf{w}_{LCMV} = \mathbf{R}^{-1}\,\mathbf{S}\left(\mathbf{S}^H\,\mathbf{R}^{-1}\,\mathbf{S}\right)^{-1}\mathbf{g}, \tag{4.10}$$

described in detail in [18] [16]. There is a risk with using a minimum variance algorithm: Since the algorithm tries to minimize the total power, it can in some cases also suppress the SOIs. Especially if the estimated directions to the SOIs are slightly wrong, and if the SOIs are strong compared to the interference, this can happen.

The *a priori* information is the directions to the SOIs which must be estimated. The satellite positions are known and with the estimates of the position and attitude for the vehicle from the navigation filter these directions can be estimated. Hence, the minimum variance algorithm is a good choice. The risk for suppressing the SOIs by mistake is small, since the power of the GPS signals is low compared to both interfering signals and the thermal noise.

4.4 Beamforming Strategies

When using the most suitable algorithm, Linear Constraint Minimum Variance (LCMV), the gain constraints can be set in different ways. It is appropriate to set gain constraints in directions to SOIs, if they are known. If the directions to interfering sources are also known, zero gain constraints can be set in those directions. However, this will probably not improve the result, since the algorithm automatically minimizes the incoming power outside the constrained directions. Instead the algorithm becomes more sensitive for errors in the constrained direction of the interfering source.

Three different types of constraints will be compared in simulations. These three are described below.

4.4.1 Single LCMV Beamformer If the constraints in (4.9) are set for all visible satellites, one weight vector $\mathbf{w}$ is calculated that fulfills all these constraints, see Figure 4.5(a).

$$\mathbf{S}^H\mathbf{w} = \mathbf{g}. \tag{4.11}$$

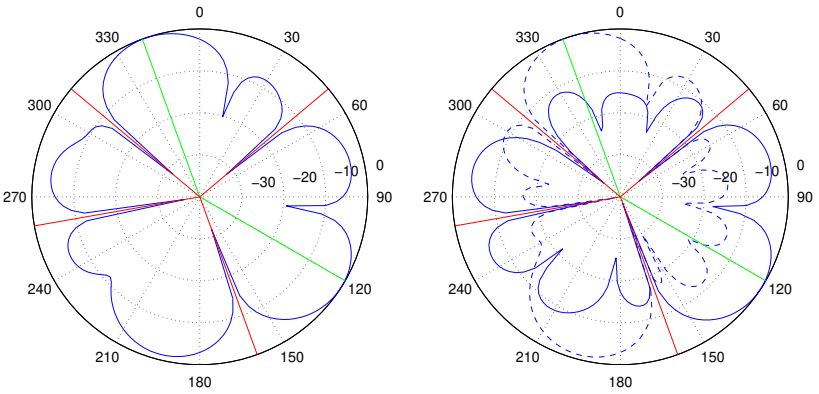
If the received power from all SOIs are the same (which is a reasonable assumption for GPS if all satellites are at approximately the same distance), also all SNRs will be the same. If one interfering source happens to be close to the direction where a constraint is set, the algorithm will not be able to suppress it, and all the SNRs for all SOIs will be degraded. This is the drawback with the single beamformer approach.

Another possible problem with the single beamformer approach is in case of a large number of SOIs. If assuming that there are L SOIs present, resulting in that L constraints are set, there will be only $(N - L)$ degrees of freedom left to form nulls for suppressing interfering signals.

4.4.2 Multiple LCMV Beamformers The constraint in (4.9) can also be set in only one direction

$$\mathbf{s}^H\mathbf{w} = g, \tag{4.12}$$

where $\mathbf{s}$ is now a steering vector. With several beamformers processing the same data in parallel, but with different constraints (4.12), it is possible to form several beam patterns at the same time, where each pattern is optimized for one certain satellite, see Figure 4.5(b). This requires that a weight vector is calculated individually for each SOI, which can be computationally costly in case of many satellites. If this is possible, and the set of constraints $\mathbf{s}$ is correct, the method gives an optimization of the SNR for all satellites, since only one constraint is set for each beamformer. This is shown in [19]. Note that the covariance matrix $\mathbf{R}$ is the same for all constraints (and so, also the inverse $\mathbf{R}^{-1}$, which is used in (4.10)) at a fixed time k.



(a) Single beamformer. One single LCMV beamformer is used to form one beam pattern that takes all constraints into account.

(b) Multiple beamformers. Two beamformers are running in parallel, each of them forming a beam pattern adapted for one of the constraints.

Figure 4.5: Sample gain patterns for single and multiple LCMV beamforming algorithms i decibel (the radial axis) for different azimuth angles. The patterns are for a static scenario with two SOIs at 120° and 340°. Four SNOIs are present at 50°, 160°, 260° and 310°.

4.4.3 Unconstrained Power Minimization If no constraints in (4.9) are set, the algorithm only tries to minimize the incoming power. This will not give a good result, since the trivial solution, which gives a minimum power, is $\mathbf{w} = \mathbf{0}$. But this will definitely not give an increased SNR. To avoid this solution, a help constraint has to be set, e.g. the constraint

$$\mathbf{s}^H \mathbf{w} = 1, \tag{4.13}$$

with $\mathbf{s} = [1, 0, \ldots, 0]^T$, is sufficient [16]. This means that the weight for the reference element should be equal to 1. One degree of freedom is spent on the help constraint, so there are $(N - 1)$ more nulls to use for suppressing interferers.

When using an unconstrained power minimization algorithm there is usually a great risk of suppressing also the SOIs. However, the GPS case is an exception since the power of the satellite signals are far below the noise floor (approximately 20 dB below).

4.5 Practical Problems and Improvements

4.5.1 High Dynamics For highly dynamic vehicles, the signal environment can change rapidly. Especially changes in the attitude give large variations in directions to the SOIs. Then it is necessary to recalculate the weights more often. But every change in the weight vector $\mathbf{w}$ gives an abrupt phase shift in the signal delivered to the GPS receiver [20]. This could lead to tracking loop problems. It is left as a future problem to be investigated how fast the weights must be updated.

4.5.2 Large Number of Interfering Signals Another problem is if there is a higher number of interfering signals than the degrees of freedom for the beamformer. This is a problem especially for the single beamformer LCMV, where many degrees of freedom are spent on forming main beams towards satellites. The only solution is to build arrays with a larger number of elements, or to avoid areas where one can expect a large number of interfering sources.

4.5.3 Errors in Steering Vector In a case when the GPS receiver loses track on all satellites due to jamming, only the INS information is available. The estimation errors in position and attitude grows large over time. This has the effect that the beamformer is forced to use uncertain *a priori* information. If the *a priori* information is considerably wrong, the beam patterns will be misdirected and the GPS receiver will have problem to find the satellite signals again. It is therefore suitable to use a method which does not require any *a priori* information. The easiest strategy is to use the array for reception of signals, but without performing any beamforming. However, this will not give any improvement of the SNR. Another, probably better, strategy is to use the Power minimization algorithm described in Section 4.4.3, which works without any direction information (no constraints). The drawback is the risk of suppressing some of the satellite signals, but that is not a serious problem as long as the SNR is increased for some other satellites.

4.5.4 Bad Directivity for Low Elevations A beamforming antenna array gives best performance, in terms of narrow beams, in the direction perpendicular to the array plane. For directions close to the plane, i.e. low elevations, the beams get wider. This is explained by the effective area of the array. For high elevations the total area of the array is directed towards the source. For lower elevation angles the effective area decreases. The effective area depends on $\sin\theta$, where θ is the elevation angle.

4.5.5 Estimating the Array Correlation Matrix The array correlation matrix, defined in (4.8), describes the spatial correlation of the incoming signals for all antenna elements. In practice the $N \times N$ covariance matrix $\mathbf{R}(k)$ must be estimated from the received data, with a finite number, M, of samples

$$\hat{\mathbf{R}}(k) = \frac{1}{M} \sum_{j=k}^{k+M-1} \mathbf{u}(j)\mathbf{u}^H(j). \tag{4.14}$$

The more number of samples M used, the better the estimate $\hat{\mathbf{R}}(k)$ gets.

Effects on the resulting beam pattern when few samples are used in the estimation of $\mathbf{R}$ are high sidelobes and a distorted main beam. A way to avoid these problems is *diagonal loading* [21] [18]. Diagonal loading means that a diagonal matrix is added to the estimated array correlation matrix $\hat{\mathbf{R}}$. Often the diagonal matrix is chosen as a function of the smallest eigenvalue of $\hat{\mathbf{R}}$. The resulting matrix is

$$\hat{\mathbf{R}}_{DL}(k) = \hat{\mathbf{R}}(k) + \alpha\,\lambda_{min}\,\mathbf{I}, \tag{4.15}$$

where λ_{min} is the smallest eigenvalue of $\hat{\mathbf{R}}$, α is the *load factor* and $\mathbf{I}$ is an identity matrix. The diagonal loading increases all eigenvalues of the covariance matrix the same amount. More about diagonal loading can be found in [21] or [2].

5. Linearization and Discretization of Nonlinear Systems

The Kalman filter is a recursive state estimation algorithm for linear systems. No further explanation of the Kalman filter is given here, but can be found in e.g. [5] or [22].

The Kalman filter theory requires linear systems, but also with non-linear systems the Kalman filter could be interesting to use. The system must also be expressed in time-discrete form for the filter to be implemented in a computer. In this chapter methods for linearizing and discretizing a non-linear time-continuous system will be described.

5.1 Discretization of a Linear System

Linear systems can not be directly implemented in discrete environments.The discretization of a continuous time system is discribed in thid section.

Assume that we have a linear time-continuous and possibly time-variant system

$$\begin{aligned} \dot{\mathbf{z}}(t) &= \mathbf{F}(t)\,\mathbf{z}(t) + \mathbf{G}(t)\,\mathbf{w}(t) \\ \mathbf{y}(t) &= \mathbf{H}(t)\,\mathbf{z}(t) + \mathbf{v}(t) \end{aligned} \tag{5.1}$$

and want to transform it to a discrete-time form

$$\begin{aligned} \mathbf{z}_k &= \mathbf{\Phi}(t_k, t_{k-1})\,\mathbf{z}_{k-1} + \mathbf{u}(k) \\ \mathbf{y}_k &= \mathbf{H}_k\,\mathbf{z}_k + \mathbf{v}_k, \end{aligned} \tag{5.2}$$

where k is a discrete time index and $\mathbf{\Phi}$ is a *state transition matrix* that will be described later.

5.1.1 System Equation We start with transforming the first equation in the continuous-time system (5.1), the *system equation*, to fit into the first of the equations in the discrete-time system (5.2), the *state transition equation*. $\mathbf{F}(t)$ is the dynamics matrix and $\mathbf{G}(t)$ is the transition matrix for the noise. Both $\mathbf{F}(t)$ and $\mathbf{G}(t)$, as well as the system state $\mathbf{z}(t)$ and the noise vector $\mathbf{w}(t)$, can be time-dependent. The noise vector is a continuous white Gaussian noise vector process. At each time t, the noise vector, $\mathbf{w}(t)$, and system states, $\mathbf{z}(t)$, are assumed to be independent.

Linear, first-order differential equations can be solved by taking the sum of the homogeneous solution and any particular solution determined by *the method of variation of parameters*, described in [4]. The solution to the non-homogeneous system equation in (5.1), with boundary conditions at t' given by $\mathbf{z}(t')$, can then be written as:

$$\mathbf{z}(t) = \mathbf{\Phi}(t, t')\mathbf{z}(t') + \int_{t'}^{t} \mathbf{\Phi}(t, \xi)\mathbf{G}(\xi)\mathbf{w}(\xi)d\xi, \tag{5.3}$$

where the state transition matrix $\mathbf{\Phi}(\cdot,\cdot)$, for a fixed initial value t', satisfies the differential equation:

$$\frac{\partial}{\partial t}\mathbf{\Phi}(t,t') = \mathbf{F}(t)\mathbf{\Phi}(t,t'), \tag{5.4}$$

with initial condition

$$\mathbf{\Phi}(t',t') = \mathbf{I}. \tag{5.5}$$

In other words, the state transition matrix $\mathbf{\Phi}(t,t')$, transforms the state from its value at time t', to the value at an arbitrary time t. For sufficiently small time intervals, $\Delta t = t - t'$, the time-continuous dynamics matrix $\mathbf{F}(t)$ is approximately constant during the interval, which is constant. With this approximation the solution to (5.4) can be written

$$\mathbf{\Phi}(t,t') \approx e^{\mathbf{F}(t')\Delta t}. \tag{5.6}$$

The right-hand side of (5.6) can be expanded as an infinite series. Truncating it to the first order gives

$$\mathbf{\Phi}(t,t') \approx \mathbf{I} + \mathbf{F}(t')\Delta t. \tag{5.7}$$

As mentioned above, $\mathbf{\Phi}(t,t')$ can be seen as a transition, taking the state vector from an initial time t', to another time, t. By identifying the initial time as $t' = t_{k-1}$ and the arbitrary time $t = t_k$ the discrete system can be written as

$$\mathbf{z}_k = \mathbf{\Phi}(t_k, t_{k-1})\mathbf{z}_{k-1} + \mathbf{u}_k, \tag{5.8}$$

where

$$\mathbf{u}_k = \int_{t_{k-1}}^{t_k} \mathbf{\Phi}(t_k,\xi)\mathbf{G}(\xi)\mathbf{w}(\xi)d\xi. \tag{5.9}$$

The fact that the integral is a linear operator and that $E[\mathbf{w}(t)] = \mathbf{0}$ for all t, implies that the expected value for $\mathbf{u}_k$ is also zero. Hence,

$$E[\mathbf{u}_k] = \mathbf{0}. \tag{5.10}$$

If the time intervals for which the integral in (5.9) is evaluated are non-overlapping, the samples $\mathbf{u}_k$ for different values of k become uncorrelated. Therefore, $\mathbf{u}_k$ is a discrete white Gaussian noise sequence.

The covariance matrix for a process $\mathbf{p}_k$ with zero mean is generally given by $E[\mathbf{p}_k\mathbf{p}_k^T]$, and therefore the covariance matrix for $\mathbf{u}_k$ can be written as:

$$E[\mathbf{u}_k\mathbf{u}_k^T] = \int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{t_k} \mathbf{\Phi}(t_k,\vartheta)\mathbf{G}(\vartheta)E[\mathbf{w}(\vartheta)\mathbf{w}(\zeta)^T]\mathbf{G}^T(\zeta)\mathbf{\Phi}^T(t_k,\zeta)d\vartheta d\zeta. \tag{5.11}$$

The covariance of a scalar time-continuous white noise process $n(t)$ with zero mean is

$$E[n(\zeta)n(\zeta')] = \sigma_n^2\,\delta(\zeta - \zeta'), \tag{5.12}$$

where σ_n^2 is the constant power spectral density of the process and $\delta(t)$ is the Dirac delta. For the vector $\mathbf{w}(t)$ of independent white noise processes this implies that

$$E[\mathbf{w}(\zeta)\mathbf{w}^T(\zeta')] = \mathbf{Q}\,\delta(\zeta - \zeta'), \tag{5.13}$$

where the diagonal matrix $\mathbf{Q}$ contains the power spectral densities of the processes in $\mathbf{w}(t)$. Using (5.13) simplifies the covariance equation (5.11)

$$E[\mathbf{u}_k\mathbf{u}_k^T] = \int_{t_{k-1}}^{t_k} \mathbf{\Phi}(t_k,\vartheta)\mathbf{G}(\vartheta)\mathbf{Q}\mathbf{G}(\vartheta)^T\mathbf{\Phi}^T(t_k,\vartheta)d\vartheta. \tag{5.14}$$

For short periods $\Delta t = t_k - t_{k-1}$, it is possible to simplify this integral. $\mathbf{G}(\vartheta)$ can then be considered constant, $\mathbf{G}_k$, during that period and $\mathbf{\Phi}(t_k, t_{k-1})$ can be approximated with an identity matrix $\mathbf{I}$. By applying these approximations, (5.14) can be written

$$\mathbf{\Theta}_k \equiv E[\mathbf{u}_k \mathbf{u}_k^T] \approx \mathbf{G}_k \, \mathbf{Q} \, \mathbf{G}_k^T \, \Delta t. \tag{5.15}$$

5.1.2 Measurement Equation The second equation of the system in (5.1), the measurement equation, is more straightforward to discretize. In the equation, the matrix $\mathbf{H}(t)$ connects the states $\mathbf{z}(t)$ with the measurements $\mathbf{y}(t)$. A vector $\mathbf{v}(t)$ with white Gaussian measurement noise is added.

The equation can simply be discretized by evaluating it at times $t = t_k$, where k is a discrete index. The noise process $\mathbf{v}(t)$ can be redefined in discrete time.

This fulfills the discretization of the system equation in (5.1).

5.2 Using Kalman Filters with Nonlinear Systems

In a situation as the one we have here, where two systems are to be integrated, a *complementary Kalman Filter* is a good choice. The complementary filter uses differences of measurements and estimates of the same quantity as observations. The benefit with this is that the filter handles error states instead of full states. The error states are expected to be small which makes it possible to use a linearized system, which, in turn, makes the Kalman filter applicable.

Two types of linearizations will be performed, the *Linearized Kalman Filter* (LKF) and the *Extended Kalman Filter* (EKF). The linearization will be applied to the navigation equations, derived in Chapter 2.

Let $\mathbf{z}(t)$ be the full navigation state (containing attitude, velocity and position) and let $\boldsymbol{\omega}(t)$ and $\mathbf{a}(t)$ represent the gyro and accelerometer outputs, respectively. If $\boldsymbol{\omega}(t)$ and $\mathbf{a}(t)$ are the ideal (error free) IMU signals, the error free navigation state $\mathbf{z}(t)$ will be a solution to the navigation equation. With these notations, the system of navigation equations can be written as

$$\begin{aligned} \dot{\mathbf{z}}(t) &= \mathbf{f}\left(\mathbf{z}(t), \boldsymbol{\omega}(t), \mathbf{a}(t)\right) \\ \mathbf{y}(t) &= \mathbf{h}\left(\mathbf{z}(t)\right) + \mathbf{v}(t). \end{aligned} \tag{5.16}$$

Here $\mathbf{f}(\cdot)$ and $\mathbf{h}(\cdot)$ are non-linear functions, $\mathbf{y}(t)$ is an observation of the error free navigation state $\mathbf{z}(t)$ and $\mathbf{v}(t)$ is white Gaussian noise.

5.2.1 Linearized Kalman Filter In the linearized Kalman Filter (LKF), it is assumed that an approximate, or nominal, trajectory $\tilde{\mathbf{z}}(t)$ is available. In the algorithm, $\tilde{\mathbf{z}}(t)$ is used as a trajectory differing $\delta\mathbf{z}(t)$ from the true trajectory $\mathbf{z}(t)$.

Now, let $\tilde{\boldsymbol{\omega}}(t)$ and $\tilde{\mathbf{a}}(t)$ be the real gyro and accelerometer signals:

$$\tilde{\boldsymbol{\omega}} = \boldsymbol{\omega} + \delta\boldsymbol{\omega} \tag{5.17}$$

$$\tilde{\mathbf{a}} = \mathbf{a} + \delta\mathbf{a}, \tag{5.18}$$

where $\delta\boldsymbol{\omega}$ and $\delta\mathbf{a}$ are the errors affecting the gyros and accelerometers, e.g. biases and noise. The corresponding solution to the navigation equations is denoted $\tilde{\mathbf{z}}$. Now we have

$$\dot{\tilde{\mathbf{z}}} = \mathbf{f}\left(\tilde{\mathbf{z}}, \tilde{\boldsymbol{\omega}}, \tilde{\mathbf{a}}\right). \tag{5.19}$$

The errors $\delta\mathbf{z}$ in the navigation states are defined according to

$$\tilde{\mathbf{z}} = \mathbf{z} + \delta\mathbf{z}. \tag{5.20}$$

If the error states $\delta\mathbf{z}$ and IMU-errors $\delta\boldsymbol{\omega}$ and $\delta\mathbf{a}$ are assumed to be small, linearization of the error free navigation state dynamics, $\dot{\mathbf{z}}$, gives

$$\begin{aligned}\dot{\mathbf{z}} = \dot{\tilde{\mathbf{z}}} - \dot{\delta\mathbf{z}} &= \mathbf{f}(\mathbf{z}, \boldsymbol{\omega}, \mathbf{a}) \\ &= \mathbf{f}(\tilde{\mathbf{z}} - \delta\mathbf{z},\, \tilde{\boldsymbol{\omega}} - \delta\boldsymbol{\omega},\, \tilde{\mathbf{a}} - \delta\mathbf{a}) \\ &\approx \mathbf{f}(\tilde{\mathbf{z}}, \tilde{\boldsymbol{\omega}}, \tilde{\mathbf{a}}) - \mathbf{F}(t)\,\delta\mathbf{z} \;- \mathbf{G}_1(t)\,\delta\boldsymbol{\omega} \;- \mathbf{G}_2(t)\,\delta\mathbf{a},\end{aligned} \tag{5.21}$$

with

$$\mathbf{F}(t) \equiv \left[\frac{\partial \mathbf{f}}{\partial \mathbf{z}}\right]_{\tilde{\mathbf{z}}(t),\tilde{\boldsymbol{\omega}}(t),\tilde{\mathbf{a}}(t)} \tag{5.22}$$

$$\mathbf{G}_1(t) \equiv \left[\frac{\partial \mathbf{f}}{\partial \boldsymbol{\omega}}\right]_{\tilde{\mathbf{z}}(t),\tilde{\boldsymbol{\omega}}(t),\tilde{\mathbf{a}}(t)} \tag{5.23}$$

$$\mathbf{G}_2(t) \equiv \left[\frac{\partial \mathbf{f}}{\partial \mathbf{a}}\right]_{\tilde{\mathbf{z}}(t),\tilde{\boldsymbol{\omega}}(t),\tilde{\mathbf{a}}(t)}. \tag{5.24}$$

Using the INS solution (5.19) in (5.21) and rearranging approximately gives

$$\dot{\delta\mathbf{z}} = \mathbf{F}(t)\,\delta\mathbf{z} + \mathbf{G}_1(t)\,\delta\boldsymbol{\omega} + \mathbf{G}_2(t)\,\delta\mathbf{a}. \tag{5.25}$$

This is the INS error equation. The approximate solution $\tilde{\mathbf{z}}$ works as a nominal solution, around which the linearization takes place.

Next, define the $\delta\mathbf{y}$ to be the difference between a calculated observation based on the nominal solution $\tilde{\mathbf{z}}$ and the real observation $\mathbf{y}$:

$$\delta\mathbf{y} \equiv\; \mathbf{h}(\tilde{\mathbf{z}}) - \mathbf{y}. \tag{5.26}$$

Using the second equation of (5.16) for the real observation gives

$$\delta\mathbf{y} = \mathbf{h}(\tilde{\mathbf{z}}) - \mathbf{h}(\tilde{\mathbf{z}} - \delta\mathbf{z}) - \mathbf{v}, \tag{5.27}$$

which, if $\delta\mathbf{z}$ is small, can be approximated with

$$\delta\mathbf{y} = \mathbf{H}(t) \cdot \delta\mathbf{z} - \mathbf{v}, \tag{5.28}$$

where

$$\mathbf{H}(t) \equiv \left[\frac{\partial \mathbf{h}}{\partial \mathbf{z}}\right]_{\tilde{\mathbf{z}}(t)}. \tag{5.29}$$

Hence, the differenced observation $\delta\mathbf{y}$ can be seen as an observation of the error state $\delta\mathbf{z}$. This is the property of a complementary filter, as previously mentioned. The two equations, (5.25) and (5.28), together form a linear state space system:

$$\begin{aligned}\dot{\delta\mathbf{z}} &= \mathbf{F}(t) \cdot \delta\mathbf{z} + \mathbf{G}(t)\,\delta\mathbf{w} \\ \delta\mathbf{y} &= \mathbf{H}(t) \cdot \delta\mathbf{z} - \mathbf{v},\end{aligned} \tag{5.30}$$

where

$$\mathbf{G}(t) = \begin{bmatrix} \mathbf{G}_1(t) & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{G}_2(t) \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{5.31}$$

$$\delta\mathbf{w} = [\,\delta\boldsymbol{\omega}^T \;\delta\mathbf{a}^T]^T. \tag{5.32}$$

The structure of the linearized system in (5.30) is similar to the general system form for a linear system in (5.1). The discretization procedure described in Section 5.1 can

therefore be applied to the system in (5.30). Using the notation from the discretized system in (5.2) results in the following linearized discrete-time system:

$$\begin{aligned} \delta \mathbf{z}_k &= \mathbf{\Phi}(t_k, t_{k-1}) \cdot \delta \mathbf{z}_{k-1} + \mathbf{G}_k \mathbf{w}_k \\ \delta \mathbf{y}_k &= \mathbf{H}_k \cdot \delta \mathbf{z}_k - \mathbf{v}_k, \end{aligned} \tag{5.33}$$

where $\delta \mathbf{z}_k$ is the error state, $\mathbf{w}_k$ and $\mathbf{G}_k$ are the discretized process noise and process noise matrix, respectively, $\mathbf{H}_k$ is the observation matrix and $\mathbf{v}_k$ is a white Gaussian measurement noise, all at the time $t = t_k$. The system in (5.33) is now linear and discrete, and the Kalman filter theory can be applied.

Prediction In the prediction step there are no measurements available. To calculate the values of the covariances $\mathbf{P}_{k|k-1}$ and predicted state estimates $\widehat{\delta \mathbf{z}}_{k|k-1}$ at time t_k, the system model is used for propagation:

$$\mathbf{P}_{k|k-1} = \mathbf{\Phi}(t_k, t_{k-1}) \mathbf{P}_{k-1} \mathbf{\Phi}^T(t_k, t_{k-1}) + \mathbf{\Theta}_k \tag{5.34}$$

$$\widehat{\delta \mathbf{z}}_{k|k-1} = \mathbf{\Phi}(t_k, t_{k-1}) \widehat{\delta \mathbf{z}}_{k-1}, \tag{5.35}$$

with $\mathbf{\Theta}_k$ defined in (5.15). Here $\mathbf{P}_{k-1}$ and $\widehat{\delta \mathbf{z}}_{k-1}$ are the covariance and state estimate, respectively, at time t_{k-1}.

Kalman Gain The Kalman gain $\mathbf{K}_k$ reflects the optimal trade-off between trusting the measurements (which gives noisy output) and trusting the propagated state estimate from (5.35) (which gives bad tracking possibility).

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k \right)^{-1}, \tag{5.36}$$

where $\mathbf{R}_k$ is the covariance matrix for $\mathbf{v}_k$.

Measurement Update When an observation is available, the Kalman filter has more information and therefore the previously predicted states are no longer the optimal ones. The updated covariances and states are then:

$$\mathbf{P}_k = \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k \right) \mathbf{P}_{k|k-1} \tag{5.37}$$

$$\widehat{\delta \mathbf{z}}_k = \widehat{\delta \mathbf{z}}_{k|k-1} + \mathbf{K}_k \left(\delta \mathbf{y}_k - \mathbf{H}_k \widehat{\delta \mathbf{z}}_{k|k-1} \right). \tag{5.38}$$

In the LKF, the states are estimates of the difference between the nominal trajectory and the true (unknown) one. An estimate $\hat{\mathbf{z}}_k$ of the full state is

$$\hat{\mathbf{z}}_k = \tilde{\mathbf{z}}_k - \widehat{\delta \mathbf{z}}_k. \tag{5.39}$$

One risk with using LKF is that the linearization may not always be valid. As can be seen in the block diagram in Figure 5.1, the LKF processes the data from the INS *before* they are corrected with the estimated error $\widehat{\delta \mathbf{z}}$. If $\delta \mathbf{z}$ is not small this means that the true solution differs much from the nominal one, and the linearization that is performed about the nominal trajectory $\tilde{\mathbf{z}}$ may not be valid.

5.2.2 Extended Kalman Filter

The Extended Kalman Filter (EKF) is similar to the LKF, but with a few important differences. One is that the linearization is performed around a trajectory estimated by the filter, not a pre-computed nominal one as in the LKF. Another difference is that the EKF will estimate the full states, not error states. It is around the filter estimate $\hat{\mathbf{z}}_{k|k-1}$ that the linearization takes place. This has the effect that the linearization is always valid.

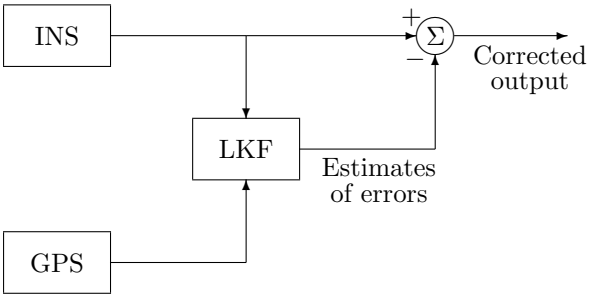


Figure 5.1: The principle of the LKF algorithm, a feed-forward filter.

Since the filter handles full states, nonlinear equations can be used in the prediction of the states. Hence, a numerical integration of the navigation equations can be used to propagate the states. The navigation equation can be written on the form

$$\dot{\mathbf{z}} = \mathbf{f}(\mathbf{z}, \boldsymbol{\omega}, \mathbf{a}), \tag{5.40}$$

where $\mathbf{z}$ is the states and $\boldsymbol{\omega}, \mathbf{a}$ are the gyro and accelerometer output, respectively.

Hence, a prediction $\tilde{\mathbf{z}}_k$ of the full state at time $t = t_k$ is given by the result of integrating the navigation equation from $t = t_{k-1}$ to $t = t_k$. The initial value at time $t = t_{k-1}$ is the current filter estimate $\hat{\mathbf{z}}_{k-1}$. The term $\tilde{\mathbf{z}}_k$ is the one-step-ahead prediction of the full state that an unaided INS would give.

Since $\tilde{\mathbf{z}}_k$ is the full state prediction, the navigation error $\delta\mathbf{z}_k$ can be defined as the difference between $\tilde{\mathbf{z}}_k$ and the true states $\mathbf{z}_k$:

$$\delta\mathbf{z}_k = \tilde{\mathbf{z}}_k - \mathbf{z}_k. \tag{5.41}$$

Hence, we can write an estimate of the full state as

$$\hat{\mathbf{z}}_k = \tilde{\mathbf{z}}_k - \widehat{\delta\mathbf{z}}_k, \tag{5.42}$$

where $\widehat{\delta\mathbf{z}}_k$ is an estimate of the navigation error $\delta\mathbf{z}_k$. Further, a prediction of $\mathbf{z}$, i.e. $\hat{\mathbf{z}}_{k|k-1}$ can be written as

$$\hat{\mathbf{z}}_{k|k-1} = \tilde{\mathbf{z}}_k - \widehat{\delta\mathbf{z}}_{k|k-1}. \tag{5.43}$$

Now, we want to write the measurement update equation

$$\widehat{\delta\mathbf{z}}_k = \widehat{\delta\mathbf{z}}_{k|k-1} + \mathbf{K}_k \left(\delta\mathbf{y}_k - \mathbf{H}_k \widehat{\delta\mathbf{z}}_{k|k-1} \right), \tag{5.44}$$

in terms of full states. The innovation term, i.e. $\delta\mathbf{y}_k - \mathbf{H}_k\widehat{\delta\mathbf{z}}_{k|k-1}$, can be rewritten using the definition of $\delta\mathbf{y}$ from (5.26):

$$\begin{aligned}
\delta\mathbf{y}_k - \mathbf{H}_k\widehat{\delta\mathbf{z}}_{k|k-1} &= \mathbf{h}(\tilde{\mathbf{z}}_k) - \mathbf{y}_k - \mathbf{H}_k\widehat{\delta\mathbf{z}}_{k|k-1} \\
&= \underbrace{\mathbf{h}(\tilde{\mathbf{z}}_k) - \mathbf{H}_k\widehat{\delta\mathbf{z}}_{k|k-1}}_{\approx \mathbf{h}(\tilde{\mathbf{z}}_k - \widehat{\delta\mathbf{z}}_{k|k-1})} - \mathbf{y}_k \\
&= \mathbf{h}(\tilde{\mathbf{z}}_k - \widehat{\delta\mathbf{z}}_{k|k-1}) - \mathbf{y}_k \\
&= \mathbf{h}(\hat{\mathbf{z}}_{k|k-1}) - \mathbf{y}_k.
\end{aligned} \tag{5.45}$$

Note that the definition of $\mathbf{H}_k$ now is

$$\mathbf{H}_k \equiv \left[\frac{\partial \mathbf{h}}{\partial \mathbf{z}} \right]_{\hat{\mathbf{z}}_{k|k-1}}. \tag{5.46}$$

Using (5.45) the measurement update equation for the error states (5.44) can be written

$$\widehat{\delta \mathbf{z}}_k = \widehat{\delta \mathbf{z}}_{k|k-1} + \mathbf{K}_k \left(\mathbf{h}(\hat{\mathbf{z}}_{k|k-1}) - \mathbf{y}_k \right). \tag{5.47}$$

If $\tilde{\mathbf{z}}_k$ is subtracted from both sides of (5.47), and using (5.43) we get

$$\hat{\mathbf{z}}_k = \hat{\mathbf{z}}_{k|k-1} + \mathbf{K}_k \left(\mathbf{y}_k - \mathbf{h}(\hat{\mathbf{z}}_{k|k-1}) \right), \tag{5.48}$$

which is the measurement update equation for the full states that will be used in the EKF.

Prediction The prediction of the covariance has not changed from the LKF algorithm:

$$\mathbf{P}_{k|k-1} = \mathbf{\Phi}(t_k, t_{k-1}) \mathbf{P}_{k-1} \mathbf{\Phi}^T(t_k, t_{k-1}) + \mathbf{\Theta}_k \tag{5.49}$$

The state prediction is given by integrating the navigation equations

$$\dot{\mathbf{z}} = \mathbf{f}(\mathbf{z}, \boldsymbol{\omega}, \mathbf{a}) \tag{5.50}$$

at time $t = t_k$, with the initial value $\mathbf{z} = \hat{\mathbf{z}}_{k-1}$ at time $t = t_{k-1}$.

Kalman Gain The Kalman gain is computed identically as in the LKF

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k \right)^{-1} \tag{5.51}$$

but with the matrix $\mathbf{H}_k$ defined in (5.46).

Measurement Update The previously determined update equation for the full states is:

$$\hat{\mathbf{z}}_k = \hat{\mathbf{z}}_{k|k-1} + \mathbf{K}_k \left(\mathbf{y}_k - \mathbf{h}(\hat{\mathbf{z}}_{k|k-1}) \right), \tag{5.52}$$

while the covariance update has not changed:

$$\mathbf{P}_k = \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k \right) \mathbf{P}_{k|k-1}. \tag{5.53}$$

Note that the matrix $\mathbf{H}_k$ now depends on the current estimate.

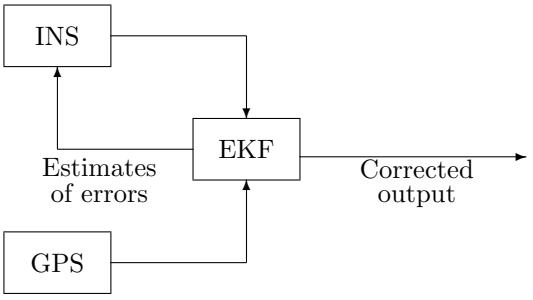


Figure 5.2: The principle of EKF, a feed-back filter.

The fact that the linearized equations depend on the previous measurements, makes the EKF difficult to analyze. There is no guarantee that the EKF performs better than the LKF in every realization, but empirically it does. Especially, the EKF is more sensitive for errors in the initial values, compared to the LKF, since initial values effect the future linearization [5]. If the initial guess is bad there is a risk that the trajectory used to linearize about in the next step is worse than the initial one, leading to an even bigger error in the next step, and so forth.

In the block diagram for the EKF in Figure 5.2, it is illustrated that the EKF processes data *after* the correction with the estimated errors. This has the effect that the linearization is always valid, since it is performed around the current state estimate.

If the covariance matrix $\mathbf{P}$ is not positive semi-definite there is a major risk for the state estimates to diverge from the correct ones. This is a problem for both the LKF and the EKF.

It should be noted that the covariance matrix $\mathbf{P}$ is computed independently of the state estimates and the observations (except that the state estimates determine the point of linearization in the EKF). Comparing the variances in the diagonal of $\mathbf{P}$ with the estimates could give information about how well the filter reflects the real navigation. If the model is different from the real world this will be obvious by comparing the variances and the estimates. Since the covariances are independent of observations, it will not be able to adapt unmodelled changes as well as the state estimates will do.

6. Implementation of Simulation Environment

As mentioned in section 1.2.1, INS and GPS have complementary characteristics. GPS pseudoranges are available at a relatively low rate (typically 1 Hz) and can give measurements with a bounded error. An INS can give a high output rate but with an unbounded error growth. Further, the INS works better under high dynamics since the oscillator in the GPS receiver is sensitive to dynamic stress. In Table 6.1 the characteristics of INS and GPS are compared.

	INS	GPS
Measurement rate	High	Low
Error growth	Unbounded	Bounded
Dynamic properties	Good	Bad
Jamming susceptibility	Very low	High

Table 6.1: Characteristics of INS and GPS

Since the INS system alone has a drift due to accelerometer and gyro biases, an important part of integrating the GPS receiver is to estimate these biases. This can only be done when there is sufficient GPS satellite information. The bias estimates can be used to reduce the errors when INS has to be used alone, for instance in a tunnel or in case of intentional (or unintentional) jamming of the GPS signals.

Another advantage with the INS/GPS integration is that the INS can support the tracking loops in the GPS. A feedback from the INS then tells the GPS about the velocities of the vehicle, which can be used to calculate the expected line-of-sight Doppler. With this knowledge the bandwidths of the tracking loops can be more narrow which leads to a smaller wideband noise power, which in turn gives a better jamming performance of the GPS. Both systems take advantage of each other, the INS uses the bounded error of the GPS and the GPS can use the accurate velocity estimate from the INS to aid the tracking loops.

6.1 Integration of GPS and INS measurements

In this thesis the integration of the INS and GPS data will be done by means of a Kalman filter. All equations will be expressed in an earth-fixed earth-centered frame (e-frame), using a Cartesian coordinate system.

The following definition of an error δz will be used:

$$\delta z = \hat{z} - z, \tag{6.1}$$

is used, where $\hat{z}$ is the estimate of the quantity z. This definition is important and will be used consistently through this thesis.

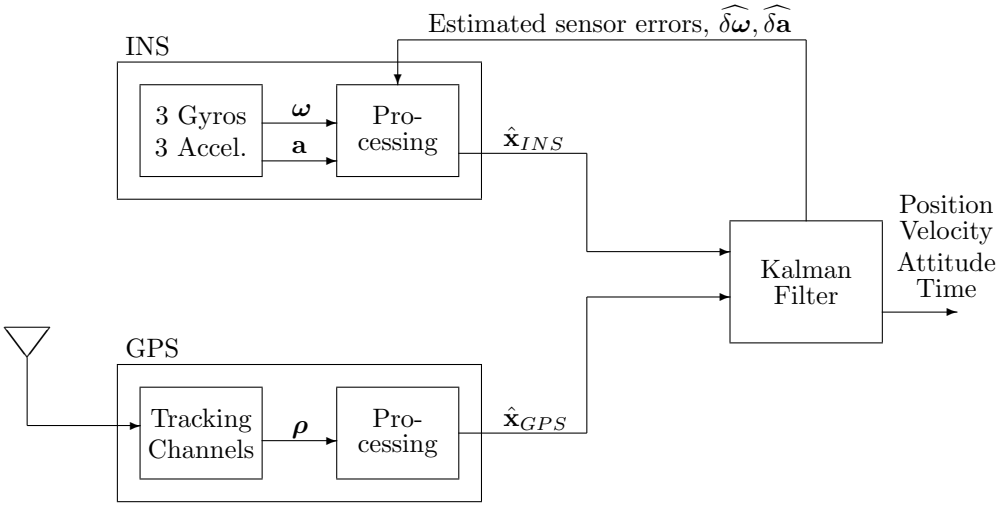


Figure 6.1: Loosely coupled INS/GPS (Decentralized integration). In the decentralized case, measurements from different sensors are processed by separate filters and the integration filter only has access to these filter outputs.

6.1.1 Coupling Approaches A tightly coupled filter will be used to integrate GPS and INS measurements. This means that the raw measurements from both the IMU and the GPS are used in the integration filter (see Figure 6.2). The opposite is a loosely coupled filter, where the raw measurements have been processed by another filter and the integration filter only has access to the estimates and their covariances (see Figure 6.1).

The cascade coupling of filters in the loosely coupled filter gives a problem. Often the second filter has access only to the variances of the states and not the cross correlation between them. This lack of information makes it impossible to design an optimal integration filter. Another problem is that positions from GPS receivers are often filtered to get a less noisy output. This has the effect that measurement errors become correlated over time.

More important is that a centralized integration filter can use GPS measurements even if less than four pseudoranges are available, which is not the case with a decentralized integration.

A loosely coupled filter is also implemented for comparing the performance with the tight coupling. The principles of the loosely and tightly coupled filters are shown in the block diagrams in Figures 6.1 and 6.2.

6.1.2 Kalman Filter Types There are different possible types of Kalman filters. In this application it is suitable to use a *complementary filter*. A complementary filter uses the difference between two estimates or measurements of the same quantity as an observation. When using an extended Kalman filter this is done by comparing the expected pseudoranges from the position predicted by the filter with the actual measured pseudoranges. Hence, the observations $\delta\mathbf{y}$, can be written

$$\begin{aligned} \delta\mathbf{y} &= \hat{\boldsymbol{\rho}} - \boldsymbol{\rho} \\ &= \mathbf{h}(\hat{\mathbf{x}}, c\,\widehat{\delta t}) - \boldsymbol{\rho}. \end{aligned} \tag{6.2}$$

where $\hat{\mathbf{x}}$ is the current position estimate, $c\,\widehat{\delta t}$ is a clock offset estimate and $\boldsymbol{\rho}$ is the vector of pseudoranges. The non-linear function $\mathbf{h}(\cdot,\cdot)$ describes how the pseudorange

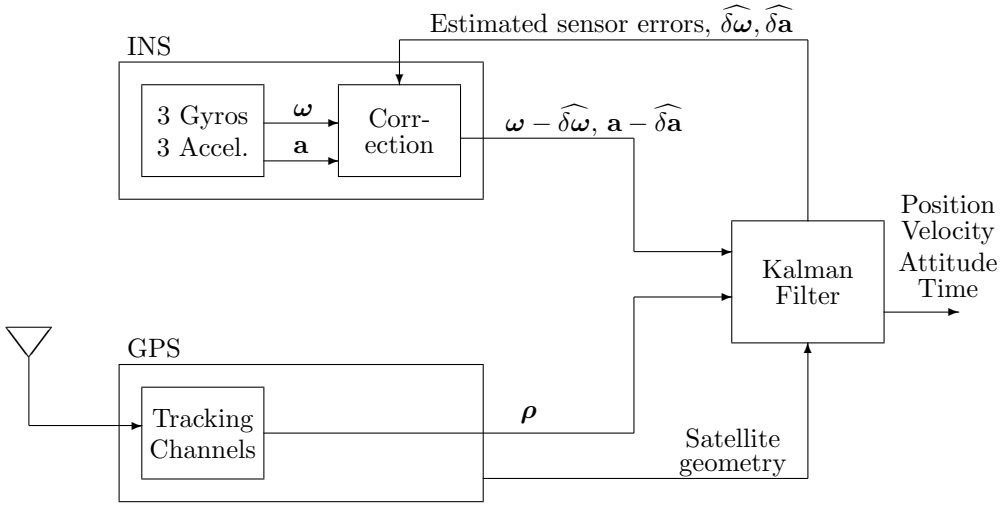


Figure 6.2: Tightly coupled INS/GPS (Centralized integration) In the tight integration the integration filter has access to both the raw inertial sensor outputs ($\boldsymbol{\omega}$ and $\mathbf{a}$) and the pseudorange measurements $\boldsymbol{\rho}$ to form the estimates and their covariances.

observations vector $\boldsymbol{\rho}$ depends on the current states. It depends on the user position $\mathbf{x}$, the clock offset $c\,\delta t$ and a Gaussian noise $\mathbf{v}$

$$\boldsymbol{\rho} = \mathbf{h}(\mathbf{x}, c\,\delta t) = \mathbf{r}(\mathbf{x}) + c\,\delta t + \mathbf{v}, \tag{6.3}$$

where $\mathbf{r}(\mathbf{x}) = [\,\|\mathbf{x}_1 - \mathbf{x}\|, \|\mathbf{x}_2 - \mathbf{x}\|, \ldots, \|\mathbf{x}_m - \mathbf{x}\|]^T$ are the geometrical distances to the m visible satellites. This implies that the observation $\delta\mathbf{y}$ in (6.2) can be written as

$$\begin{aligned} \delta\mathbf{y} &= (\hat{\mathbf{r}} + c\,\delta t) - (\mathbf{r} + c\,\delta t + \mathbf{v}) \\ &= \hat{\mathbf{r}} - \mathbf{r} - \mathbf{v} \\ &= \delta\mathbf{r} - \mathbf{v}, \end{aligned} \tag{6.4}$$

where $\hat{\mathbf{r}}$ is the range that the position estimated by the filter would give and $\delta\mathbf{r}$ is the range error. Hence, the complementary observations $\delta\mathbf{y}$ are observations of the pseudorange prediction errors $\delta\mathbf{r}$ affected by a noise $\mathbf{v}$. The observations $\delta\mathbf{y}$ of the pseudorange prediction errors can in turn be used to calculate the navigation error $\delta\mathbf{x}$, in the same way as set of pseudoranges are used to calculate a position $\mathbf{x}$.

The Kalman filter is a linear filter and can thus only be used for linear systems. Two types of linearization approaches were addressed in Chapter 5. The Linearized Kalman filter (LKF) is linearized around a pre-defined, nominal trajectory, while the Extended Kalman filter (EKF) is linearized around the current trajectory estimated by the filter. If the linearization should be valid, it is important that the linearization point does not differ too much from the true point.

The best choice of filter type, LKF or EKF, depends on the application. In a scenario where there is a nominal trajectory that makes sense, e.g. for a UAV (Unmanned Aerial Vehicle) that is pre-programmed to follow a certain path, an LKF is the simplest choice. The true trajectory of the UAV will probably not differ very much from the expected and the linearization is always valid. In other scenarios where there is no pre-defined trajectory the EKF is to prefer. The EKF is also the most suitable if errors that should be fed back to the INS or the GPS are estimated. Since biases for the

gyros and accelerometers in the INS should be estimated in this implementation, the EKF is the best choice in this case. When the biases are fed back, the prediction in the Kalman filter performs much better, also during periods when no GPS measurements are available.

Both Kalman filter types are implemented, but only the EKF is used in simulations.

6.2 Sensor Data Generation

6.2.1 Data for the INS In an INS, signals from the gyros and accelerometers in an IMU, are used for calculating a trajectory by means of the navigation equations described in Chapter 2.

In the simulation environment this is done in the opposite order: First a trajectory is defined. The ideal gyro and accelerometer outputs, $\boldsymbol{\omega}$ and $\mathbf{a}$, that this trajectory would give are extracted. Next, errors are introduced to the ideal IMU signals to form the erroneous gyro and accelerometer signals, $\tilde{\boldsymbol{\omega}}$ and $\tilde{\mathbf{a}}$:

$$\tilde{\boldsymbol{\omega}} = \boldsymbol{\omega} + \mathbf{d}_\omega + \delta\boldsymbol{\omega} \tag{6.5}$$

$$\tilde{\mathbf{a}} = \mathbf{a} + \mathbf{b}_a + \delta\mathbf{a}, \tag{6.6}$$

where $\mathbf{d}_\omega$ and $\mathbf{b}_a$ are the constant gyro and accelerometer biases, $\delta\boldsymbol{\omega}$ and $\delta\mathbf{a}$ are two noise processes. The variances of the two noise processes $\delta\boldsymbol{\omega}$ and $\delta\mathbf{a}$ are time-invariant and have the same variance in all three components

$$\mathrm{var}(\delta\boldsymbol{\omega}) = [\, q_\omega \;\; q_\omega \;\; q_\omega \,]^T \tag{6.7}$$

$$\mathrm{var}(\delta\mathbf{a}) = [\, q_a \;\; q_a \;\; q_a \,]^T. \tag{6.8}$$

6.2.2 GPS Pseudoranges In the simulation environment the GPS will provide the integrated system with pseudorange measurements. The model for the GPS receiver is rather simple. The geometric distance r between the receiver and the satellite is calculated and delays and disturbances are added according to

$$\rho = r + c\,(\delta t - \delta t_s) + \delta r_{ion} + \delta r_{trop} + \delta r_{SA} + \delta r_{MP} + v \tag{6.9}$$

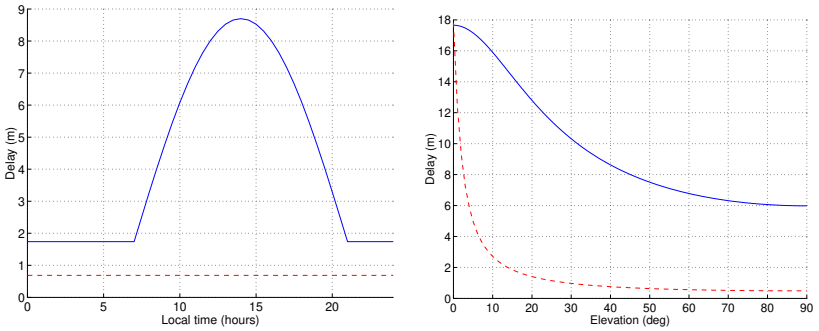
to get a pseudorange ρ. The same notations as in (3.8) are used.

Here, clock offsets (δt and δt_s) are set as constants. The atmospheric disturbance (δr_{ion} and δr_{trop}) are modelled as described in [9]. Examples of these atmospheric errors are shown in Figure 6.3.

The SA and multipath-error (δr_{SA} and δr_{MP}) generation is taken from [23]. However, these two error terms are not used in the simulations, i.e. they are set to zero. Remember that the intentionally introduced SA error was actually removed in May 2000. The last term, v, is a Gaussian measurement noise. This is the only uncorrelated (white) error, all other terms have a time-correlation.

Besides the pseudorange (which means the *code* pseudorange) there are other measurable quantities, e.g. the *phase pseudorange*. However, if the phase pseudorange should be used, the integer ambiguity (described in Section 3.2.1) has to be resolved. A search algorithm for finding the integer ambiguity has been implemented, but mainly for test purpose and is not used in simulations. The most straight-forward way to handle the problem with integer ambiguities is to assume that they are resolved. Then the phase pseudorange, which usually has an order of magnitude smaller noise compared to the code pseudorange, can be used.

Doppler measurements are not implemented. The reason for this is that the position measurements (or equivalently, pseudorange measurements) are enough to bound the



(a) Atmospheric delays for different time of day at elevation angle $\theta = 45°$.

(b) Atmospheric delays for different elevation angles θ at 12.00, local time.

Figure 6.3: Atmospheric delays for the code pseudorange, L1 band, 1575.42 MHz. The upper solid curves are the ionospheric delays and the bottom dashed curves are the tropospheric delays.

error from the INS.

Differential GPS (DGPS), which uses pseudorange corrections from a basestation to decrease the effect of atmospherical and satellite orbit errors, is the only differential method that is implemented. Double differential measurements, which eliminates the receiver clock offset is not implemented because this clock offset can be estimated from the measurements instead.

Due to the simple GPS receiver model, where realistic errors are introduced to the true range to form the pseudoranges (see (6.9)), no tracking loops are implemented. Hence, velocity aided tracking loops can not be treated.

6.3 Filter Implementation

Since the filter will be a complementary filter, the states will be error states. The system, after linearization, can be written in continuous time on the state-space form

$$\begin{aligned} \dot{\mathbf{z}}(t) &= \mathbf{F}(t)\,\mathbf{z}(t) + \mathbf{G}(t)\,\mathbf{w}(t) \\ \mathbf{y}(t) &= \mathbf{H}(t)\,\mathbf{z}(t) + \mathbf{v}(t). \end{aligned} \tag{6.10}$$

However, the implementation of the system will take place in a discrete environment, hence a time-discrete Kalman filter will be used. The discretization process was addressed in Section 5.1

6.3.1 Process Noise Modelling The last term in the system equation in (6.10), $\mathbf{G}(t)\,\mathbf{w}(t)$, is the system noise. In the derivation of the Kalman filter it is assumed that $\mathbf{w}(t)$ is a white gaussian process, caused by noise in gyros and accelerometers. Suppose that $\mathbf{w}(t)$ can be written

$$\mathbf{w}(t) = [\delta\boldsymbol{\omega}^T \;\; \delta\mathbf{a}^T\,]^T \tag{6.11}$$

where $\delta\boldsymbol{\omega}$ is the gyro noise process and $\delta\mathbf{a}$ is the accelerometer noise process. The matrix $\mathbf{G}(t)$ is

$$\mathbf{G}(t) = \begin{bmatrix} -\mathbf{C}_b^e(t) & \mathbf{0}_{3\times 3} \\ \mathbf{0}_{3\times 3} & \mathbf{C}_b^e(t) \\ \mathbf{0} & \mathbf{0} \\ \vdots & \vdots \end{bmatrix}, \tag{6.12}$$

which is block-diagonal with blocks $-\mathbf{C}_b^e(t)$, $\mathbf{C}_b^e(t)$. The term $\mathbf{C}_b^e(t)$ is a time-dependent direction cosine matrix. The transpose of a direction cosine matrix is the same as the inverse, i.e. $(\mathbf{C}_b^e(t))^T = (\mathbf{C}_b^e(t))^{-1}$. According to (5.15) the covariance matrix $\mathbf{\Theta}_k$ for the discretized process noise $\mathbf{u}_k$ is

$$\mathbf{\Theta}_k \approx \mathbf{G}_k \, \mathbf{Q} \, \mathbf{G}_k^T \, \Delta t. \tag{6.13}$$

Using the property in (6.7) and (6.8), the time-invariant covariance matrix $\mathbf{Q}$ can be written as

$$\mathbf{Q} = \begin{bmatrix} q_w \cdot \mathbf{I}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & q_a \cdot \mathbf{I}_3 \end{bmatrix}, \tag{6.14}$$

where q_w and q_a are the gyro and accelerometer noise variances, respectively. Since the structure of $\mathbf{Q}$ and $\mathbf{G}(t)$ are similar with only block-diagonal elements, the product in (6.13) is

$$\begin{aligned} \mathbf{\Theta}_k \approx \mathbf{G}_k \mathbf{Q} \mathbf{G}_k^T \Delta t &= \begin{bmatrix} q_w \cdot (-\mathbf{C}_b^e)_k \mathbf{I}_3 (-\mathbf{C}_b^e)_k^T & \mathbf{0}_3 & \mathbf{0} & \dots \\ \mathbf{0}_3 & q_a \cdot (\mathbf{C}_b^e)_k \mathbf{I}_3 (\mathbf{C}_b^e)_k^T & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \Delta t \\ &= \begin{bmatrix} q_w \cdot \mathbf{I}_3 & \mathbf{0}_3 & \mathbf{0} & \dots \\ \mathbf{0}_3 & q_a \cdot \mathbf{I}_3 & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \Delta t \end{aligned} \tag{6.15}$$

and thus independent of time. This is one of the rewards for expressing the equations in the e-frame!

6.3.2 Filter Models The error for an unaided INS is unbounded after a long time, due to imperfections in the IMU. When the INS is aided by a GPS the error can be made bounded. Another advantage with the integration is that biases in the IMU can be estimated and compensated for. This gives a better navigation performance, also during periods with bad satellite availability, when the system has to rely on the INS alone.

Three different types of GPS measurements are used in the integration:

A. A loosely coupled filter, aided by a single GPS

B. A tightly coupled filter, aided by a single GPS

C. A tightly coupled filter, aided by differential GPS

The three models are described in this section. Before describing the three models, we start with looking at the implementation of an unaided INS.

Unaided INS Since the integration will be performed by means of a complementary filter, it will handle error states. States for position error $\delta\mathbf{x}$, velocity error $\delta\mathbf{v}$ and attitude error $\delta\mathbf{\Psi}$ are used. Each of these is a 3-dimensional vector (3×1), since we should not be limited to navigate on a plane surface or only in a certain direction. The error state vector for the unaided INS is denoted $\boldsymbol{\varepsilon}_{INS}$.

$$\boldsymbol{\varepsilon}_{INS} = \begin{bmatrix} \mathbf{\Psi}^T & \delta\mathbf{v}^T & \delta\mathbf{x}^T \end{bmatrix}^T \tag{6.16}$$

The linear (9×9) system matrix from (2.31) for small errors in this system is

$$\mathbf{F}_{INS} = \begin{bmatrix} -\mathbf{\Omega}_{ie}^e & \mathbf{0}_3 & \mathbf{0}_3 \\ [\mathbf{a}\times] & -2\,\mathbf{\Omega}_{ie}^e & \mathbf{\Upsilon} \\ \mathbf{0}_3 & \mathbf{I}_3 & \mathbf{0}_3 \end{bmatrix}, \tag{6.17}$$

using the same notation as in (2.31).

With an unaided INS there are no external measurements, and hence, there is no measurement equation.

Loosely coupled, single GPS receiver In the loosely coupled filter, positions are used as observations. A position can only be calculated by the GPS if four or more satellites are available. Hence, if less than four satellites are available, no observations will be passed to the filter. This is the main drawback with the loosely coupled filter.

The error state vector now reads

$$\boldsymbol{\varepsilon}_A = \left[\begin{array}{ccc|cc} \boldsymbol{\Psi}^T & \delta\mathbf{v}^T & \delta\mathbf{x}^T & \mathbf{d}^T & \mathbf{b}^T \end{array} \right]^T, \tag{6.18}$$

where $\mathbf{d}^T$ and $\mathbf{b}^T$ are the gyro and accelerometer biases, respectively. The system matrix $\mathbf{F}_A$ for the loosely coupled filter reads

$$\mathbf{F}_A = \begin{bmatrix} \mathbf{F}_{INS} & \mathbf{F}_{12} \\ \mathbf{0}_{6\times 9} & \mathbf{0}_{6\times 6} \end{bmatrix}, \tag{6.19}$$

where the matrix $\mathbf{F}_{12}$ is

$$\mathbf{F}_{12} = \begin{bmatrix} -\mathbf{C}_b^e & \mathbf{0}_{3\times 3} \\ \mathbf{0}_{3\times 3} & \mathbf{C}_b^e \\ \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3} \end{bmatrix}. \tag{6.20}$$

The observation function $\mathbf{h}_A(\cdot)$ is trivial since the position states are observed

$$\mathbf{h}_A(\mathbf{x}) = \mathbf{x}. \tag{6.21}$$

Tightly coupled, single GPS receiver In the tightly coupled integration filter, raw pseudoranges from the GPS are used as observations. This has the advantage that the integration filter gets information even when less than four satellites are available.

The tight integration require that also error states for the GPS receiver are included. This gives an error state vector

$$\boldsymbol{\varepsilon}_B = \left[\begin{array}{ccc|cc|c} \boldsymbol{\Psi}^T & \delta\mathbf{v}^T & \delta\mathbf{x}^T & \mathbf{d}^T & \mathbf{b}^T & \delta(c\,\delta t) \end{array} \right]^T, \tag{6.22}$$

where the first nine states, $\boldsymbol{\Psi}$, $\delta\mathbf{v}$, $\delta\mathbf{x}$, are the navigation errors, defined in the previous section. The next two (3×1) vectors, $\mathbf{d}$ and $\mathbf{b}$, are gyro and accelerometer biases, respectively, and $\delta(c\,\delta t)$ is the GPS receiver clock offset error.

The system matrix now reads

$$\mathbf{F}_B = \begin{bmatrix} \mathbf{F}_{INS} & \mathbf{F}_{12} & \mathbf{0}_{9\times 1} \\ \mathbf{0}_{6\times 9} & \mathbf{0}_{6\times 6} & \mathbf{0}_{6\times 1} \\ \mathbf{0}_{1\times 9} & \mathbf{0}_{1\times 6} & \mathbf{0}_{1\times 1} \end{bmatrix}, \tag{6.23}$$

with $\mathbf{F}_{INS}$ and $\mathbf{F}_{12}$ defined in (6.17) and (6.20). In the tightly coupled filter the observations are pseudoranges and the non-linear observation function $\mathbf{h}_B$ looks like

$$\mathbf{h}_B(\mathbf{x}, c\,\delta t) = \begin{bmatrix} \|\mathbf{x}_1 - \mathbf{x}\| + c\,\delta t \\ \|\mathbf{x}_2 - \mathbf{x}\| + c\,\delta t \\ \vdots \\ \|\mathbf{x}_m - \mathbf{x}\| + c\,\delta t \end{bmatrix}, \tag{6.24}$$

where $\mathbf{x}_s$ is the position of satellite s. Note that the size of $\mathbf{h}_B$ depends on how many satellites m that are visible.

Tightly coupled, DGPS with two receivers Integrating an INS with a DGPS is similar to integrating INS with a stand-alone GPS. The difference is that the measured pseudoranges are corrected with terms from a basestation. Errors that are common in the basestation and the roving receiver cancel, i.e. atmospheric errors and satellite clock errors.

The single differential corrected pseudorange vector $\boldsymbol{\rho}_{\Delta 1}$ can be written as

$$\boldsymbol{\rho}_{\Delta 1} = \boldsymbol{\rho} - \delta\boldsymbol{\rho}_{corr}, \tag{6.25}$$

where $\boldsymbol{\rho}$ is the measured pseudoranges and $\delta\boldsymbol{\rho}_{corr}$ is the correction terms received from the basestation.

6.3.3 State Modelling

In addition to the navigation states (position, velocity and attitude), states for receiver clock offset, gyro and accelerometer biases are included. These are modelled as random constants, i.e. the time derivative is zero.

6.3.4 Numerical Problems in Covariance Update

Some numerical problems can occur in the Kalman filter. The most sensitive part is the prediction and measurement update of the covariance matrix $\mathbf{P}_k$. Two requirements on $\mathbf{P}_k$ are that it should be symmetric (since it is a covariance matrix) and positive semi-definite (the positive semi-definiteness guarantees that no diagonal elements are negative). However, it can happen that some of these requirements are not fulfilled if any of the matrices used in the Kalman filter algorithm are very ill-conditioned. This can e.g. happen if there are large differences in the scaling of the states, so that the elements in $\mathbf{P}_k$ has an extremely large spread. A non-positive-definite $\mathbf{P}_k$ usually leads to divergence of the filter [5].

The symmetry of $\mathbf{P}_k$ is easily checked. It is possible to force $\mathbf{P}_k$ to be symmetric in every iteration by replacing $\mathbf{P}_k$ with

$$\mathbf{P}_k \leftarrow \frac{1}{2}\left(\mathbf{P}_k + \mathbf{P}_k^T\right) \tag{6.26}$$

Another, more sophisticated, way is to use a symmetric form in the covariance measurement update. The usual form, repeated from (5.37) and (5.53), is

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}$$

An alternative measurement update form [5] is

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T, \tag{6.27}$$

which is the sum of two symmetric matrices, the first being positive definite and the second positive semi-definite. This expression is less sensitive to numerical errors in K_k and retains both symmetry and positive semi-definiteness better than the (5.37) form [22]. This equation for the $\mathbf{P}_k$ update is referred to as the *Symmetric Joseph form*. The symmetric form does not help for numerical problems due to an ill-conditioned $\mathbf{P}$ matrix.

Another method that can remedy the problem with asymmetry, and also help with the ill-conditioning is the *Square Root algorithm* [22][24]. The idea with all types of square root algorithms is to update the square root, $\mathbf{P}^{1/2}$, of the covariance matrix. Taking the square root decreases the span between the smallest and the largest elements in the matrix. The benefit with this is that the elements can be represented more accurately with a finite numerical precision. In practice, the numerical precision is always limited. Normally, a square root algorithm can give the same accuracy with only half the number of bits to represent a number, compared to a conventional algorithm. The

price for increased numerical precision is a slightly increased computational cost. In [22] several types of square root algorithms are compared.

The symmetric Joseph form of the covariance update equation is used in the toolbox. As a comparison, a square root algorithm has been implemented and tested with the EKF. It gives better results for situations with large initial spread in the covariance matrix. This could happen if one has very little knowledge about a certain state, so the initial variance for it has to be set very high.

6.3.5 Observability The integrated INS/GPS system that has been implemented has at least 10 states, and in most cases more. Most of the time, much less than 10 satellites will be visible. It is clear that the number of states can be higher than the number of observations. The question is: Is it possible to estimate all of the states from the given observations?

In some cases it is obvious that certain states can not be estimated. States that do not affect the observations at all, not even indirectly, are impossible to estimate. Also several states that affect the observations in the same way can not be told from each other.

To determine whether all states can be estimated, the *observability* of the system can be calculated. From [25] , an *observability matrix* $\boldsymbol{\mathcal{O}}$ for a time-invariant system, can be set up as

$$\boldsymbol{\mathcal{O}}(\boldsymbol{\Phi}, \mathbf{H}) = \begin{bmatrix} \mathbf{H} \\ \mathbf{H}\boldsymbol{\Phi}^1 \\ \mathbf{H}\boldsymbol{\Phi}^2 \\ \vdots \\ \mathbf{H}\boldsymbol{\Phi}^{m-1} \end{bmatrix} \tag{6.28}$$

where m is the number of states and $\mathbf{H}$ and $\boldsymbol{\Phi}$ are the time-discrete observation and state transition matrices, respectively. The system is fully observable only if the matrix $\boldsymbol{\mathcal{O}}$ has full rank. In the integrated INS/GPS system, both the transition matrix and the observation matrix depend on time. The observability matrix will then also be time dependent and can be written as

$$\boldsymbol{\mathcal{O}}(\boldsymbol{\Phi}_k, \mathbf{H}_k)_k = \begin{bmatrix} \mathbf{H}_{k+1-m} \\ \mathbf{H}_{k+2-m}\boldsymbol{\Phi}_{(k+2-m,k+1-m)} \\ \mathbf{H}_{k+3-m}\boldsymbol{\Phi}_{k+3-m,k+2-m}\boldsymbol{\Phi}_{k+2-m,k+1-m} \\ \vdots \\ \mathbf{H}_{k-1}\boldsymbol{\Phi}_{k-1,k-2} \;\cdots\; \boldsymbol{\Phi}_{k+3-m,k+2-m}\boldsymbol{\Phi}_{k+2-m,k+1-m} \\ \mathbf{H}_k\boldsymbol{\Phi}_{k,k-1}\boldsymbol{\Phi}_{k-1,k-2} \;\cdots\; \boldsymbol{\Phi}_{k+3-m,k+2-m}\boldsymbol{\Phi}_{k+2-m,k+1-m} \end{bmatrix} \tag{6.29}$$

where $\boldsymbol{\Phi}_{k,k-1}$ is a short-hand notation for a state transition matrix $\boldsymbol{\Phi}(t_k, t_{k-1})$ that propagates the state from time $t = t_{k-1}$ to $t = t_k$, and $\mathbf{H}_k$ is the observation matrix at time k.

Calculating the rank of $\boldsymbol{\mathcal{O}}(\boldsymbol{\Phi}_k, \mathbf{H}_k)_k$ in simulations indicates that the system is fully observable if one or more satellites are available. The exception is the start of a flight, where the number of observable states grows from zero before the first measurement to m after some time. How fast the observability grows, depends on the number of available satellites.

6.4 Integration of Adaptive Beamforming Antenna

The adaptive beamforming algorithms used in the simulations need the directions to all visible satellites expressed with azimuth ϕ and elevation θ, defined in an antenna coordinate frame according to Figure 4.4. Since coordinates in the body frame are

given in the order (*forward*, *right*, *down*), this means that azimuth is the angle from the forward direction and the elevation is the angle above the forward-right-plane. Estimates of ϕ and θ can be calculated from the state estimates given by the Kalman filter. The estimates for the position $\hat{\mathbf{x}}^e = (\hat{x}, \hat{y}, \hat{z})^T$ and the orientation relative to Earth $\hat{\mathbf{C}}_b^e$ are required. Also the satellite positions $\mathbf{s}$ in the e-frame must be known. The antenna is assumed to be strapped to the vehicle, which means that the antenna is fixed in the b-frame.

With satellite s at coordinates $\mathbf{x}_s^e = (x_s, y_s, z_s)^T$ in e-frame, an estimate of the vector from the vehicle to the satellite $\hat{\mathbf{r}}^b = (\hat{r}_1, \hat{r}_2, \hat{r}_3)^T$ in b-frame can written as

$$\hat{\mathbf{r}}^b = (\hat{\mathbf{C}}_b^e)^T(\mathbf{x}_s^e - \hat{\mathbf{x}}^e). \tag{6.30}$$

When the vector $\hat{\mathbf{r}}^b$ is known, the directions to the satellite can be estimated as

$$\hat{\phi} = \arctan_2(\hat{r}_2, \hat{r}_1) \tag{6.31}$$

$$\hat{\theta} = \arctan_2(-\hat{r}_3, \sqrt{\hat{r}_1^2 + \hat{r}_2^2}), \tag{6.32}$$

where $\arctan_2(\cdot, \cdot)$ is the *four-quadrant arctangent*. The minus sign before $\hat{r}_3$ in the expression for the elevation is due to the fact that the 3-axis in b-frame is defined to point downwards (see Section 2.2), but we want the elevation to be positive upwards, to be consistent with the definition of mask angle when talking about GPS.

6.4.1 Lever-arm Compensation If the IMU and the GPS antenna are spatially separated, the two instruments will sense slightly different positions and velocities [7]. If the separation is large, it should be considered whether this needs to be modelled. In this thesis no compensations for the lever-arm are made.

6.5 Error Measurements

6.5.1 Circular Error Probable Circular error probable (CEP) is a two-dimensional measure for position errors. It is defined as the circle centered at the true position in which 50 % of all position estimates end up. The CEP is related to the (2×2) covariance matrix describing the uncertainty in the two-dimensional position estimate. The CEP is often used when talking about GPS errors.

6.5.2 Root Mean Square The Root mean square (RMS) is a commonly used quantity for expressing the magnitude of errors that vary over time. It is, as the name indicate, defined to be the square root of the time average of the squares. For a discrete time series $\mathbf{z} = [\, z_1, z_2, \ldots, z_N]^T$ this is

$$RMS \equiv \sqrt{\frac{1}{N}\, \delta\mathbf{z}^T \delta\mathbf{z}}. \tag{6.33}$$

7. Simulation Results for GPS/INS Integration

The input data in the simulations are generated and therefore known, which is not the case in a real application. The advantage working with known data is that estimation errors can be calculated and the same set of data can be used repeatedly comparing different methods. The known estimation errors can be compared to the variance calculated by the Kalman filter. For a well-tuned filter, the true error and the standard deviation should correspond in size. A different behavior indicates that some unmodelled errors are present.

The simulation has been divided into four sections:

Simulations without atmospheric disturbance To be able to compare the different integration methods based on different observations, simulation without atmospherical errors are performed in Section 7.2

Simulations with atmospheric disturbance In practice, there will always be influence from the signal propagation through the atmosphere. In Section 7.3 it is tested how these unmodelled disturbances affect the performance.

Simulations with satellite outage During some periods, the GPS receiver can lose the ability to track some of the satellites. This could be caused by either blocking of the signal or other interfering signals. The results of satellite outages are tested in Section 7.4.

Comparison between different flight paths The type of flight path will affect the results. In Section 7.5 the estimation errors are compared for a low dynamic and a more high dynamic flight path.

The results are summarized in Section 7.6.

7.1 General Settings

The simulations are made with a relatively low IMU sample rate, 10 Hz. Usually this rate is much higher, $200 - 400$ Hz is common. The reason for the low rate is to speed up the simulations. Since the same rate is used in all simulation, the results are comparable anyway. The GPS pseudoranges are measured with the rate of 1 Hz.

7.1.1 Integration Methods Different integrating methods have been implemented. All methods use an extended Kalman filter (EKF) to integrate the INS and GPS data. In all integration methods gyro and accelerometer biases are estimated and used for compensation in the time propagation of the EKF which takes place between the GPS measurements. A short description is given below of the three different integration methods:

A. Loosely coupled, single receiver This method uses a loosely coupled extended Kalman filter. The observations used in a loose integration are positions. This means that an external process has solved the possibly over-determined system of equations by least squares to convert pseudoranges into position.

B. Tightly coupled, single receiver In this integration a single GPS receiver and an INS system are tightly coupled. This means that the measurements used in the Kalman filter are pseudoranges to GPS satellites. An advantage with using pseudoranges instead of positions as observations is that pseudoranges are available even if less than four satellites are in line-of-sight.

C. Tightly coupled DGPS A base station is located at known coordinates and receives information from the same satellites as a moving receiver (rover). Since the position of the base station is known, a correction term for each pseudorange can be calculated. As long as the rover can be considered to have the same atmospheric disturbances as the base station, the correction terms can be used to remove the atmospheric influence. This requires that the baseline between base station and rover must be relatively short. This is assumed to be fulfilled. The integration filter used is identical to the one used for the tightly coupled, single receiver.

7.1.2 GPS Satellite Constellation The simulations are performed in GPS week number 167 and starts at 43 200 s. This corresponds to UTC (Coordinated Universal Time) 12.00 on Sunday, November 3^{rd} 2002.

At this time there were 9 satellites visible if a 5° mask angle is used. The mask angle sets a lower limit of the angle above the horizon for the satellites that are used. The visible satellites and their constellation in the sky are illustrated in Figure 7.1. This satellite constellation gives an initial GDOP value of 2.15, increasing to about 2.25 after 10 minutes.

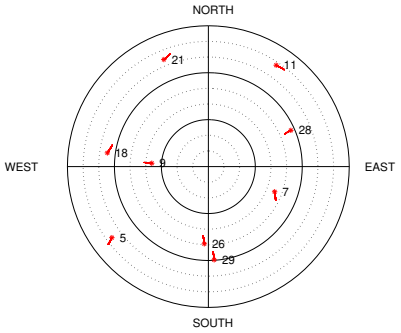


Figure 7.1: Visible satellites during the simulations. The satellites positions are drawn in polar coordinates (azimuth and elevation) from a user point of view. The outer circle corresponds to the horizon and center of the circles is zenith. No satellites below the mask angle 5° are used.

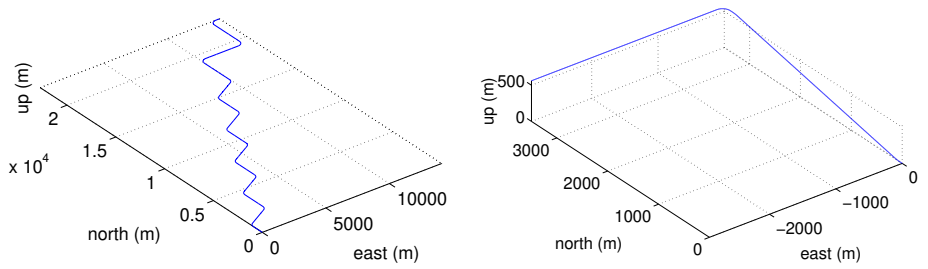
7.1.3 Flight Paths Two different flight paths, shown in Figure 7.2, have been used in the simulations. Both flight paths start in Stockholm, Sweden, with coordinates

$$\mathbf{x}^e = [\,3124205,\ 1015116,\ 5467131\,]^T \quad [\text{m}]$$

in e-frame, at zero velocity. In geodetic coordinates this corresponds to latitude 59.1692° and longitude 18.0000°.

The main difference between the two flight paths is the dynamics. The first flight path (Figure 7.2(a)) contains several turns. Each turn manoeuver consists of one roll, one pitch and finally, one roll back again. The second flight path (Figure 7.2(b)) contains only one pitch manoeuver in the beginning and one 90° turn.

The flight path in Figure 7.2(a) is a test path used by the Swedish Defence Material Administration to test the performance of flight sensors [26].



(a) The high dynamic flight path

(b) The low dynamic flight path

Figure 7.2: An illustration of the flight paths used in the simulations. Both paths last for about 11 minutes but they differ in their dynamics.

7.1.4 IMU Performance The IMU introduces errors. These are modelled as white noise and biases in the accelerometer and gyro sensors. The variances of the white noise and the values of the biases are taken from a Litton LN200 A1 data sheet and are listed in Table 7.1. This IMU is considered in [4] as a medium quality IMU, but in other books, equally performing IMUs are said to be of low quality.

Litton LN200 A1	White noise spectral density	Bias
Gyro	$0.09°\ \mathrm{h}^{-1/2}$	$10°\ \mathrm{h}^{-1}$
	$0.09°\ \mathrm{h}^{-1/2}$	$10°\ \mathrm{h}^{-1}$
	$0.09°\ \mathrm{h}^{-1/2}$	$10°\ \mathrm{h}^{-1}$
Accelerometer	$4.9050 \cdot 10^{-4}\ \mathrm{m\,s^{-2}Hz^{-1/2}}$	$2.9 \cdot 10^{-3}\ \mathrm{m\,s^{-2}}$
	$4.9050 \cdot 10^{-4}\ \mathrm{m\,s^{-2}Hz^{-1/2}}$	$2.9 \cdot 10^{-3}\ \mathrm{m\,s^{-2}}$
	$4.9050 \cdot 10^{-4}\ \mathrm{m\,s^{-2}Hz^{-1/2}}$	$2.9 \cdot 10^{-3}\ \mathrm{m\,s^{-2}}$

Table 7.1: Errors introduced on the INS sensors. Values are taken from a Litton LN200 A1 data sheet.

7.2 Simulation without Atmospheric Disturbance

In order to compare the performance of method A, the loose integration, and method B, the tight integration, a first simulation is done without any atmospherical disturbances. The flight path depicted in Figure 7.2(a) is used.

Introduced Errors In all simulations, a set of pseudoranges $\boldsymbol{\rho}$ from user to satellites are generated. In a tight integration (method B) these pseudoranges are also the observations. In the loose integration (method A) a least square calculation is performed on each observed set of pseudoranges, to achieve position coordinates as observations. This has been done to achieve comparable results in the simulation. The errors that are introduced to these pseudoranges are listed in Table 7.2.

Kalman Filter Parameters To get an optimal performance of the filter, the statistics of the errors affecting the measurements must be chosen as good as possible using data sheets and testing. Here, the errors are generated and therefore known, but in reality errors are typically unknown. A measurement error covariance matrix $\mathbf{R}$ should reflect the errors in Table 7.2. The measurements are assumed to be uncorrelated, which is reasonable for pseudoranges, with a variance $\sigma^2 = (2.0\ \mathrm{m})^2$. This value does not take the receiver clock offset into account, since it is estimated as a state and the uncertainty is affecting through the state covariance matrix $\mathbf{P}$.

Error sources	Introduced error σ	σ in estimation
Ionosphere	0 m	0 m
Troposphere	0 m	0 m
Selective Availability	0 m	0 m
Multi-path	0 m	0 m
Receiver clock offset	2 μs ($\approx$ 600 m)	0 m
Satellite clock offset	0 m	0 m
Uncorrelated noise	2.0 m	2.0 m
User Equivalent Range Error (UERE)		2.0 m

Table 7.2: Errors added to the pseudorange measurements. The values in the left column are the introduced errors while the values in the right column are used to form the measurement covariance matrix $\mathbf{R}$. The reason for not including the receiver clock offset when forming the UERE is that it is estimated as a state and affects the filter through the state covariance matrix $\mathbf{P}$.

The initial value for the receiver clock offset in the state covariance matrix is set to 1000 m. All initial values for $\mathbf{P}$ are presented in Table 7.3.

Attitude	Velocity	Position	Gyro bias	Acc bias	Clock offset
$5 \cdot 10^{-3}$ rad	0.1 μm/s	10 m	$4.85 \cdot 10^{-5}$ rad/s	$2.9 \cdot 10^{-3}$ m/s^2	1000 m
$5 \cdot 10^{-3}$ rad	0.1 μm/s	10 m	$4.85 \cdot 10^{-5}$ rad/s	$2.9 \cdot 10^{-3}$ m/s^2	
$5 \cdot 10^{-3}$ rad	0.1 μm/s	10 m	$4.85 \cdot 10^{-5}$ rad/s	$2.9 \cdot 10^{-3}$ m/s^2	

Table 7.3: Square roots of the diagonal elements in the initial error covariance matrix $\mathbf{P}_0$. The parameters with three values belong to variables in three dimensions, e.g. position.

The system noise covariance $\mathbf{Q}$ is another tunable parameter. The values used are listed in Table 7.4. The variances for the attitude and velocity system noise reflect the gyro and accelerometer noise listed in Table 7.1. The bias and clock offset states are modelled as random constants (i.e. their time derivative is zero), and will therefore have no system noise.

Attitude	Velocity
$0.6854 \cdot 10^{-9}$ (rad/s)^2	$0.2406 \cdot 10^{-6}$ (m/s)^2
$0.6854 \cdot 10^{-9}$ (rad/s)^2	$0.2406 \cdot 10^{-6}$ (m/s)^2
$0.6854 \cdot 10^{-9}$ (rad/s)^2	$0.2406 \cdot 10^{-6}$ (m/s)^2

Table 7.4: Diagonal elements in the Kalman filter system noise matrix $\mathbf{Q}$. The noise variance for the attitude and velocity reflect the gyro and accelerometer white noise spectral densities, respectively.

We also need an initial estimate of the state vector. The start position is assumed to be known, so that attitude, velocity and position can be initialized with the correct values. The bias estimates, on the other hand (gyro, accelerometer and receiver clock offset), are all initialized to zero, which is the only reasonable guess for a bias. If the biases should have an expected value which is not zero, it should not need to be modelled as a *random* constant, but rather compensated for as a deterministic constant (possibly combined with an estimated random constant).

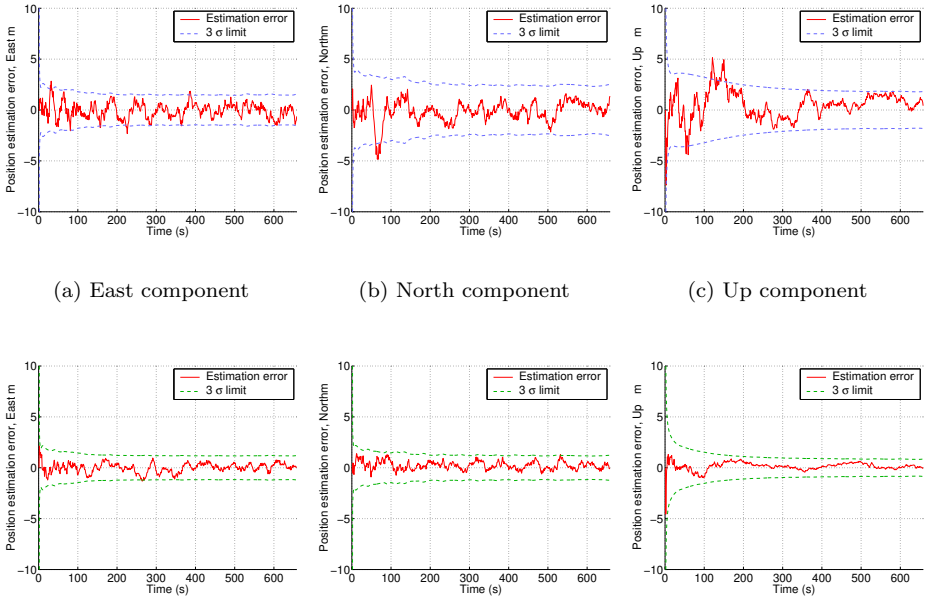
Methods used Two methods are tested and compared, the loose integration (method A) and the tight integration (method B). The filter configurations are described in Section 6.3.2. The biases in the IMU sensors (see Table 7.1) will also be estimated by

the filters, and used to compensate the IMU outputs. The compensation of the IMU outputs are depicted in Figure 6.2.

Loose integration (method A) The loose integration approach uses complementary measurements of the position as observation. The only information available to the filter is a three-dimensional position estimate and its estimated variance. The estimated position error variance is given by the variance of the pseudoranges and the square of the current DOP. The diagonal elements of the measurement error covariance matrix is set depending on the current DOP and the pseudorange errors. Hence, the filter will not be able to handle correlation between the errors in the three coordinates.

Tight integration (method B) The tightly integrated filter uses the complementary measured pseudoranges as observations. In theory, method B should give a better result than method A, since more information is available to the filter.

The results from the simulations without atmospheric disturbance are shown in Figures 7.3, 7.4, 7.5, 7.6 and 7.7. With method B the receiver clock offset must also be estimate, which is shown in Figure 7.8. Position estimation errors, $\|\hat{\mathbf{x}} - \mathbf{x}\|$, from the simulations without atmospheric disturbances are presented in Table 7.5. The values are calculated as the mean over 5 simulations with different noise realizations. The maximum position error, position error at end point and RMS position errors are compared for the two integration methods. The RMS is defined in (6.33).

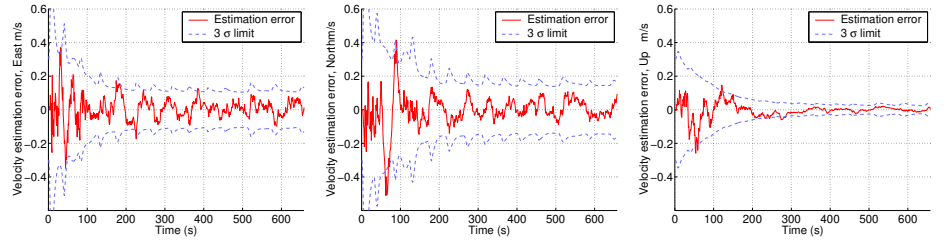


(a) East component (b) North component (c) Up component

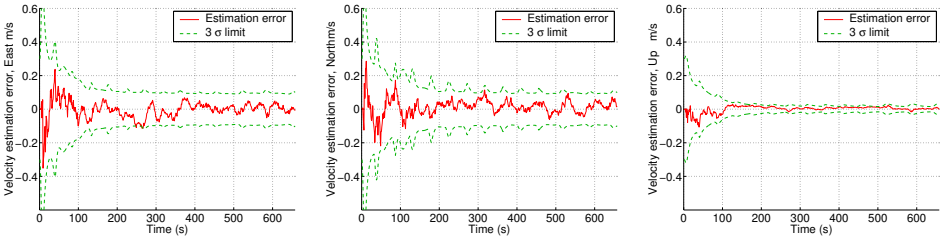
(d) East component (e) North component (f) Up component

Figure 7.3: Position estimation error for method A and method B, without atmospheric disturbance (local frame). In (a), (b), (c) method A (loose integration) was used while method B (tight integration) was used in (d), (e), (f).

Comments As expected, the tight integration (method B) gives a significantly better performance in the position estimate compared to the loose integration (method A). This can be seen in Figure 7.3. Due to the correlated measurement errors with

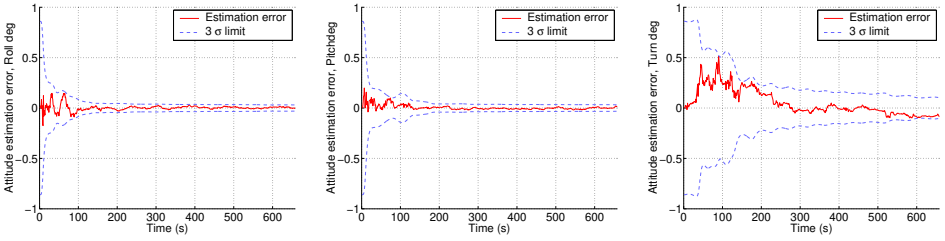


(a) East component (b) North component (c) Up component

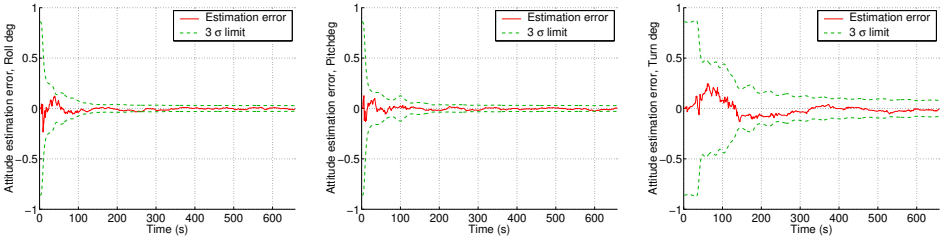


(d) East component (e) North component (f) Up component

Figure 7.4: Velocity estimation error for method A and method B, without atmospheric disturbance (local frame). In (a), (b), (c) method A (loose integration) was used while method B (tight integration) was used in (d), (e), (f).

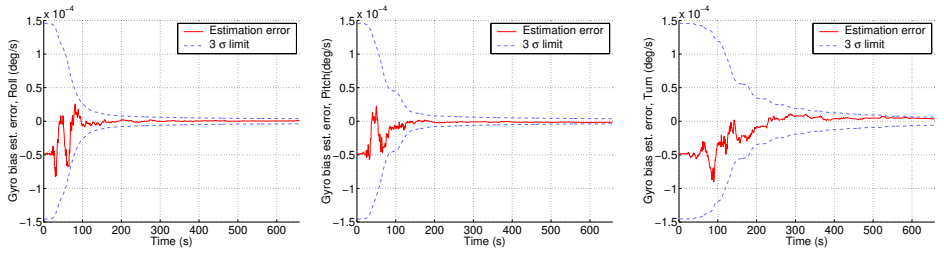


(a) Roll (b) Pitch (c) Yaw

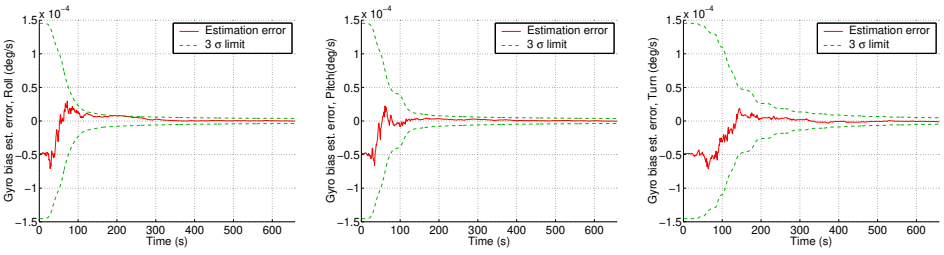


(d) Roll (e) Pitch (f) Yaw

Figure 7.5: Attitude estimation error for method A and method B, without atmospheric disturbance (local frame). In (a), (b), (c) method A (loose integration) was used while method B (tight integration) was used in (d), (e), (f).

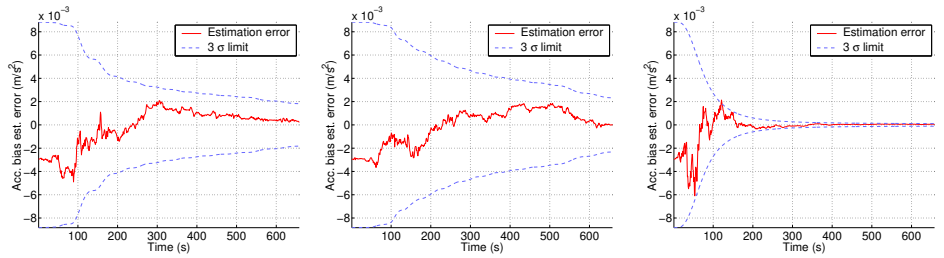


(a) Roll gyro (b) Pitch gyro (c) Yaw gyro

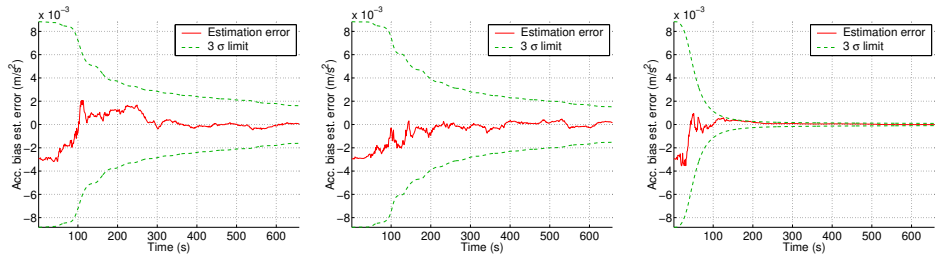


(d) Roll gyro (e) Pitch gyro (f) Yaw gyro

Figure 7.6: Gyro bias estimation error for method A and method B, without atmospheric disturbance (body frame). In (a), (b), (c) method A (loose integration) was used while method B (tight integration) was used in (d), (e), (f).



(a) Forward accelerometer (b) Right accelerometer (c) Down accelerometer



(d) Forward accelerometer (e) Right accelerometer (f) Down accelerometer

Figure 7.7: Accelerometer bias estimation error for method A and method B, without atmospheric disturbance (body frame). In (a), (b), (c) method A (loose integration) was used while method B (tight integration) was used in (d), (e), (f).

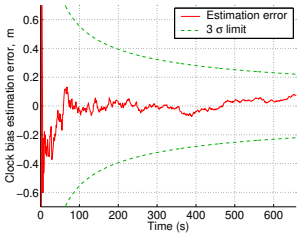


Figure 7.8: Receiver clock offset estimation error for method B (tight integration), without atmospheric disturbance.

method A, which are not modelled in the filter, the indicated covariance seems slightly mismatched compared to the real position error.

Method	Position error (m)		
	Maximum	**At end point**	**RMS**
Loosely integrated INS/GPS (A)	8.20 m	1.08 m	1.09 m
Tightly integrated INS/GPS (B)	4.18 m	0.6038 m	0.45 m

Table 7.5: Position estimation errors without atmospherical disturbance. The numbers are averaged over 5 realizations.

7.3 Simulation with Atmospheric Disturbance

The behavior of the filter when pseudoranges are affected by atmospherical disturbances is tested. A DGPS simulation (method C) is tested and compared to the tight integration (method B) that was used in the previous section. The integration filter used in method B is not designed to handle these extra non-modelled errors.

The same flight path as in the simulations without atmospheric disturbances is used (see Figure 7.2(a)).

Introduced Errors The difference from the previous simulation is that atmospherical delays are introduced. Note that the ionospheric and tropospheric delays are not actually Gaussian distributed and definitely not uncorrelated in time. The introduced delays depend both on the elevation angle and the time of the day. Their behavior are depicted in Figure 6.3 on page 51. The models for these atmospheric phenomena are taken from [9]. The errors are listed in Table 7.6.

Error sources	Introduced error σ	σ in estimation
Ionosphere	$\approx$ 6.5 m	6.5 m
Troposphere	$\approx$ 1 m	1 m
Selective Availability	0	0
Multi-path	0	0
Receiver clock offset	2 μs ($\approx$ 600 m)	0
Satellite clock offset	0	0
Uncorrelated noise	2.0 m	2.0 m
User Equivalent Range Error (UERE)		6.9 m

Table 7.6: Errors added to the pseudorange measurements. The values in the left column are the introduced errors while the values in the right column are used to form the measurement covariance matrix $\mathbf{R}$. The reason for not including the receiver clock offset when forming the UERE is that it is estimated as a state and affects the filter through the state covariance matrix $\mathbf{P}$.

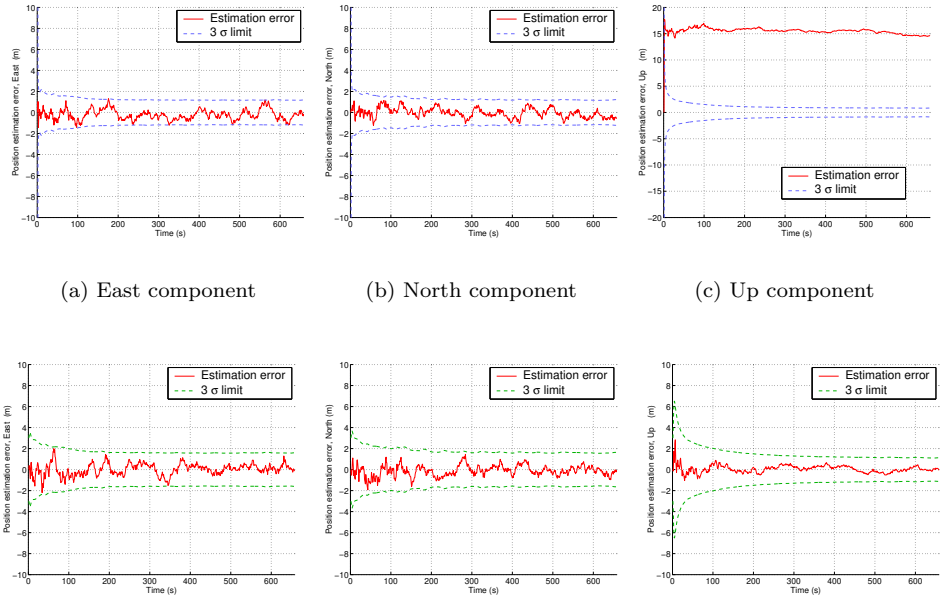
Kalman Filter Parameters The settings for the Kalman filter parameters are the same as in the previous case (see Table 7.4 on page 60).

Methods used Two methods are tested and compared:

Tight integration (B) This scenario is identical with the one tested in the previous section, except that atmospherical disturbance has been added. Since no states are added in the model to handle the extra delays, one should not expect similar performance.

DGPS (C) Using a DGPS theoretically removes all atmospheric disturbance if the baseline between the basestation and the roving receiver is short. Also satellite clock errors and disturbances in the orbits are cancelled. The only drawback is that the magnitude of the variance for the white measurement noise is twice as large, since two signals with the same noise variance are subtracted.

The results from the simulations with atmospheric disturbance are shown in Figures 7.9, 7.12, 7.10 and 7.11. Position estimation errors for the simulation with atmospheric disturbance are collected in Table 7.7. The values are calculated as the mean over 5 simulations with different noise realizations.

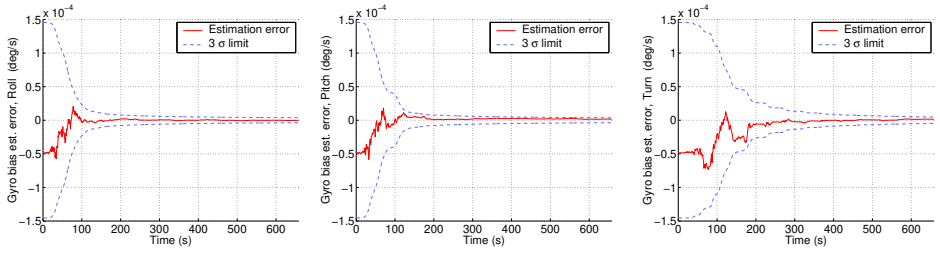


(a) East component (b) North component (c) Up component

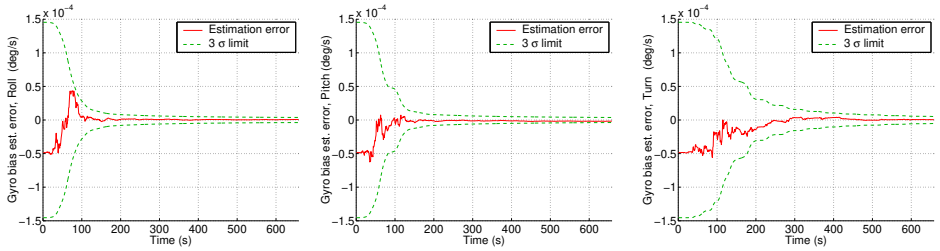
(d) East component (e) North component (f) Up component

Figure 7.9: Position estimation error for method B and method C, with atmospheric disturbance (local frame). In (a), (b), (c) method B (single receiver, tight integration) was used while method C (differential GPS, tight integration) was used in (d), (e), (f). Note that the unmodelled atmospheric disturbances when using method B, mainly effect the up-component.

Comments Differential GPS is an effective way to overcome the problem with atmospheric delays. Using a stand-alone GPS without modelling the atmospheric delays results in a bias in the estimation error. This bias mainly effects the height component of the position estimate (see Figure 7.9(c)). The reason for this is that the satellites are spread on a half-sphere from the user's point of view, which causes an unmodelled

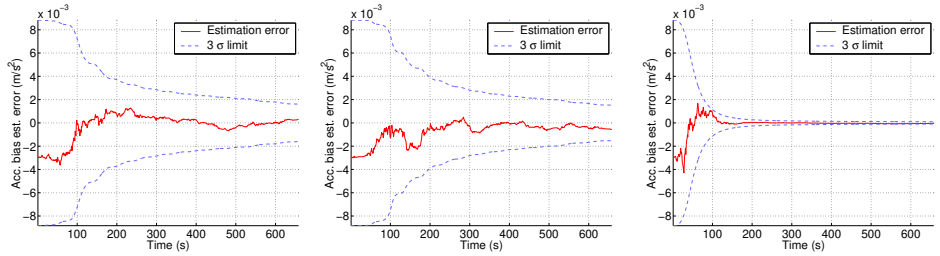


(a) Roll gyro (b) Pitch gyro (c) Yaw gyro

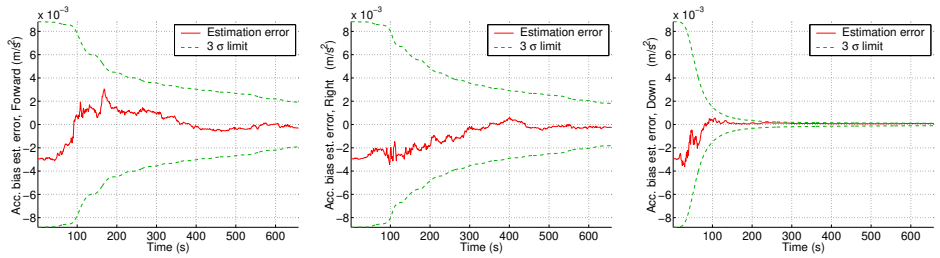


(d) Roll gyro (e) Pitch gyro (f) Yaw gyro

Figure 7.10: Gyro bias estimation error for method B and method C, with atmospheric disturbance (body frame). In (a), (b), (c) method B (single receiver, tight integration) was used while method C (differential GPS, tight integration) was used in (d), (e), (f).

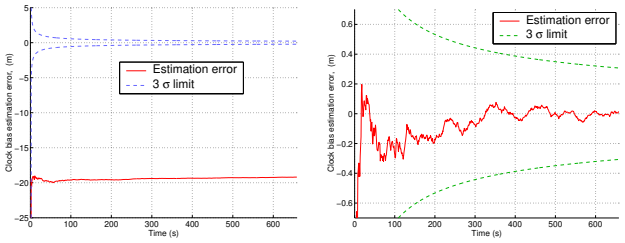


(a) Forward accelerometer (b) Right accelerometer (c) Down accelerometer



(d) Forward accelerometer (e) Right accelerometer (f) Down accelerometer

Figure 7.11: Accelerometer bias estimation error for method B and method C, with atmospheric disturbance (body frame). In (a), (b), (c) method B (single receiver, tight integration) was used while method C (differential GPS, tight integration) was used in (d), (e), (f).



(a) Method B (b) Method C

Figure 7.12: Clock offset estimation error for method B and method C, with atmospheric disturbance. In (a) method B (single receiver, tight integration) was used while method C (differential GPS, tight integration) was used in (b). The estimate in (a) is biased because the filter was not designed to handle the atmospheric delays. Note that the scales are different.

Method	Position error (m)		
	Maximum	**At end point**	**RMS**
Tightly integrated INS/DGPS	3.392 m	0.279 m	0.501 m
Tightly integrated INS/GPS	17.81 m	14.61 m	8.97 m

Table 7.7: Position estimation errors without atmospherical disturbance.

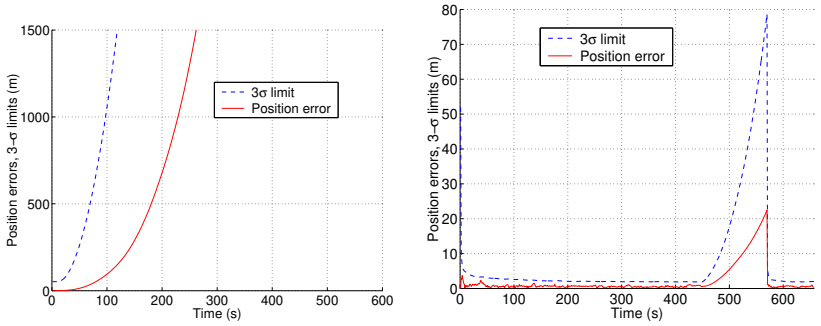
bias to affect the up-component, whereas the effect are more or less cancelled for the east and north components, if the satellite constellation is good. Also the estimate of the receiver clock offset is effected which can be seen in Figure 7.12.

7.4 Simulations with Satellite Outage

The orbits for the GPS satellites are designed in a way that there should always be at least four satellites visible. Despite this, situations with temporary satellite outages are common. The outages could be caused by intentional jamming in the GPS band or loss of line-of-sight to satellites. In urban environments it is common with very few visible satellites due to blocking of the satellite signals.

In these simulations the navigation filter is tested when the number of available satellites are reduced. The flight path depicted in Figure 7.2(a), which lasts for 660 seconds, is used. After 450 seconds a satellite outages of 120 seconds is simulated. The tight integration (method B) is used. The general settings for this simulation are the same as for the simulation in Section 7.2.

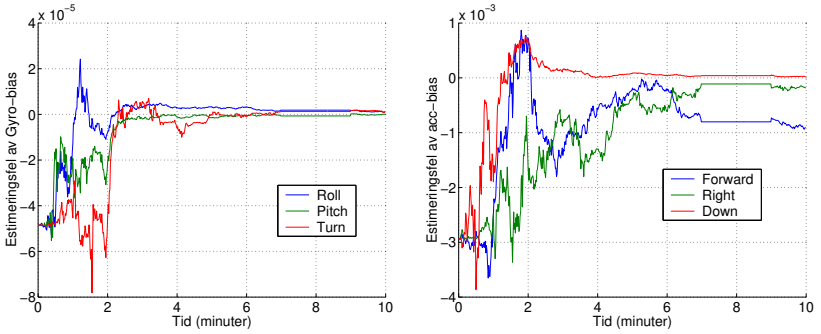
7.4.1 Bias Estimation During Satellite Outage In this simulation it is shown how systematic errors (biases) in the IMU affect the navigation performance. The biases can be estimation and compensation for in the navigation filter. Again, integration method B is used. In Figure 7.13 the performance of an unaided INS and an INS aided by a GPS are compared. It is easy to see that the navigation performance is increased if the biases are properly estimated. Note that the error grows much slower during the outage (Figure 7.13(b)) if the biases are properly estimated, even though the GPS does not get any information in this interval. In Figure 7.13(a), the stand-alone INS, the estimation error has grown to just over 150 meters after the first 120 seconds. In Figure 7.13(b), where the INS is aided by a GPS except during the outage, the estimation error is only just over 20 meters after a 120 s total satellite



(a) Unaided INS, no bias estimation

(b) INS aided by GPS with bias estimation

Figure 7.13: Position errors. Between between 450 s and 570 s a total satellite outage occur and no satellites are possible to use. In (a) the biases are neither estimated nor compensated for. This corresponds to an unaided INS or a case when the navigation filter is initialized during a long satellite outage. In (b) the navigation filter has had the chance to estimate the biases during the first 450 s. Between 450 s and 570 s no GPS information is available, but thanks to the estimated biases the error grows much slower than in (a). Note that the scales are different in (a) and (b).



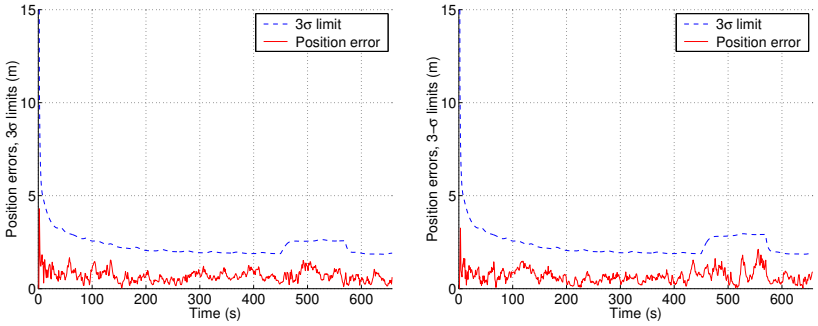
(a) Gyro bias estimation error

(b) Accelerometer bias estimation error

Figure 7.14: Bias estimation error for a flight path with a total satellite outage. Note that during the outage (450 s-570 s) the bias estimates are kept constant by the filter since no additional information is available.

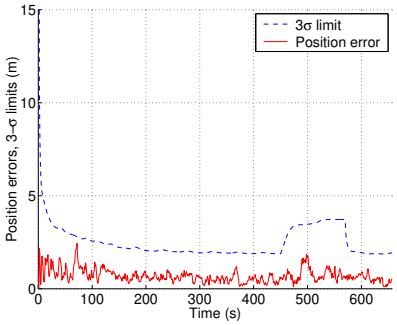
outage. The reason for this increased performance is, as previously mentioned, that constant biases in the inertial sensors can be estimated and compensated for. The bias estimation errors are depicted in Figure 7.14.

7.4.2 Position Errors for Different Number of Satellites An important difference between integration method A (loose integration) and method B (tight integration) is when the number of available satellites is reduced to less than four. Using method A, no observation at all will be presented to the filter, while with method B the measurements from the satellites in view (even when less than four) can be used as observations.

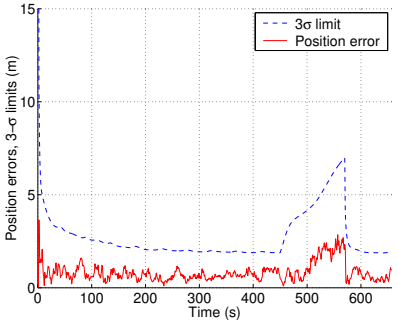


(a) 5 satellites (SV 5, SV 7, SV 9, SV 11, SV 18)

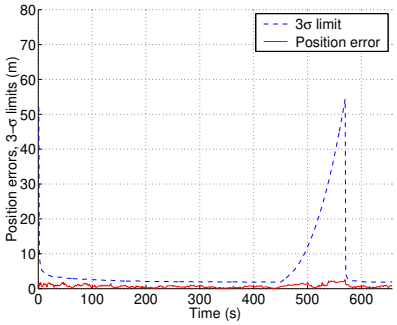
(b) 4 satellites (SV 5, SV 7, SV 9, SV 11)



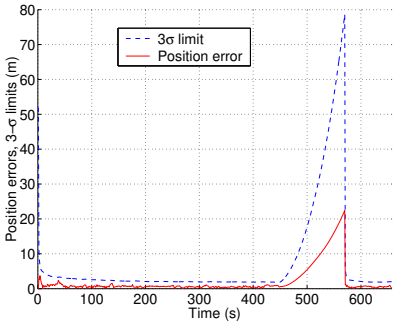
(c) 3 satellites (SV 5, SV 7, SV 9)



(d) 2 satellites (SV 5, SV 7)



(e) 1 satellite (SV 5)



(f) Total satellite outage (identical to 7.13(b))

Figure 7.15: Position errors with reduced satellite availability between 450 s and 570 s. Integration method B is used. Before 450 s and after 570 s there are 9 satellites available.

In this simulation it is tested how integration method B performs during satellites outage. Different numbers of satellites are available during the outages.

Position estimation errors during the satellite outage are shown in Figures 7.15. The differences between the six simulations are the number of available satellites during the outages, which is varied between zero and five satellites.

In Table 7.8 the changes in the variance of the position errors are listed for the satellite outage simulations.

	Number of satellites available during outage					
	5	4	3	2	1	0
Sigma at 450 s (m)	1.88	1.88	1.88	1.88	1.88	1.88
Sigma at 570 m (m)	2.58	2.91	3.72	7.00	54.4	78.7
Maximum sigma during outage (m)	2.64	2.95	3.73	7.00	54.4	78.7
Average sigma rate (m/s)	0.0058	0.0086	0.015	0.043	0.44	0.64

Table 7.8: Changes in position sigma (square root of variance) during outage.

Comments Satellite outages give larger errors in position, since less information is passed to the filter. With a tightly coupled filter satellite information can be used also if less than four GPS satellites are available. The fewer satellites that are available, the faster the estimation error and the corresponding variance grows.

In cases with only one or two satellites, the error indicated by the filter (i.e. the covariance matrix) is often larger than the real estimation error. The reason for this is that the variances indicated by the filter are slightly overestimated compared to the real estimation errors for the gyros and accelerometer biases. This can be seen e.g. in Figures 7.6 and 7.7 where the estimation error are much smaller than the $3-\sigma$ indicated by the covariance matrix in the filter.

7.5 Comparing Flight Paths

In this simulation the performance of integration method B (tight integration, single receiver) is tested for two different flight paths. The objective with this simulation is to see how the dynamics of the flight path affects the bias estimations.

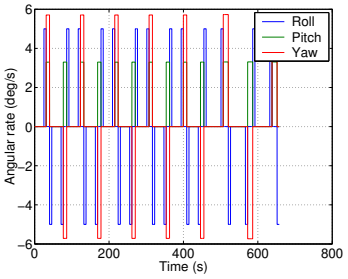
The introduced errors are the same as for the first simulations, defined in Table 7.2 and Table 7.1. The two flight paths that were used are shown in Figure 7.2. Flight path (a) is more dynamic than flight path (b). Both flight paths have the same total length. The dynamics of the flight paths are depicted in Figure 7.16.

The bias estimates for the two flights are compared in Figures 7.17 (gyro bias estimates) and in Figures 7.18 (accelerometer bias estimates).

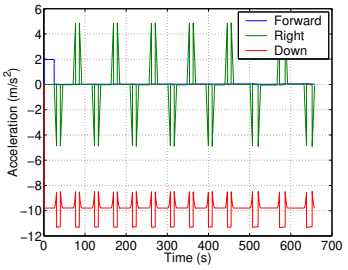
Comments The turns and rolls in the more dynamic flight path (Figure 7.2(a)) help the bias estimate to converge faster. The yaw gyro, which reflects rotations around the up-axis has a slower convergence than the other two gyros. The down accelerometer, which is affected by the gravity component most of the time, is the most easily estimated of the accelerometers.

7.6 Summary of GPS/INS Integration Simulations

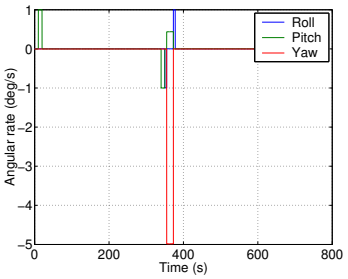
Using a tight integration, the INS can be aided by the GPS also with less than four GPS satellites available. The fewer satellites, the faster the variance estimated by the filter grows.



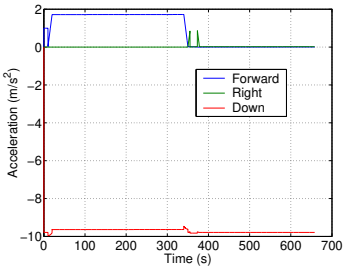
(a) Angular rate, high dynamic



(b) Acceleration, high dynamic

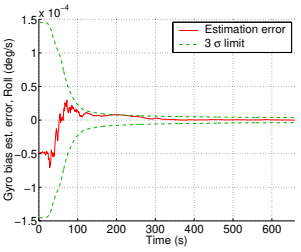


(c) Angular rate, low dynamic

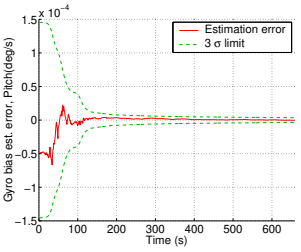


(d) Acceleration, low dynamic

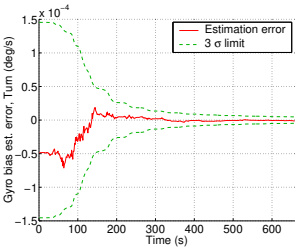
Figure 7.16: The dynamics for the two flight paths compared.



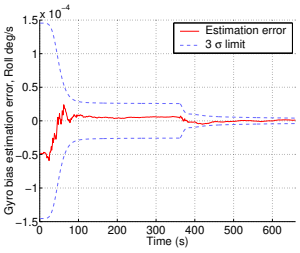
(a) Roll gyro



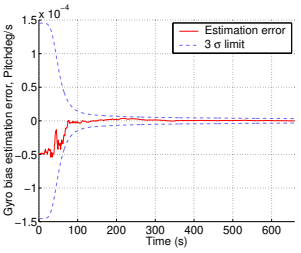
(b) Pitch gyro



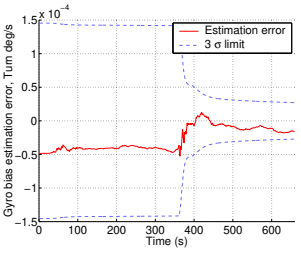
(c) Yaw gyro



(d) Roll gyro

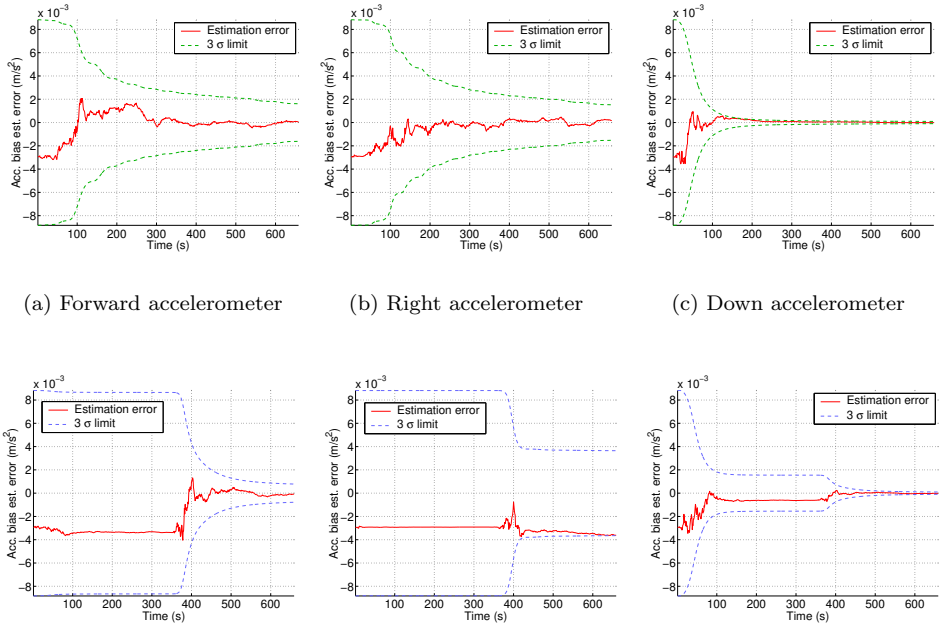


(e) Pitch gyro



(f) Yaw gyro

Figure 7.17: Gyro bias estimation errors with the high dynamic (a), (b), (c) and low dynamic flight path (d), (e), (f).



(a) Forward accelerometer (b) Right accelerometer (c) Down accelerometer

(d) Forward accelerometer (e) Right accelerometer (f) Down accelerometer

Figure 7.18: Accelerometer bias estimation errors with the high dynamic (a), (b), (c) and low dynamic flight path (d), (e), (f).

The bias estimates are noise sensitive, especially in the beginning of a flight. The flight path, and mainly the dynamics of it, is important for the performance of the bias estimation. A more high dynamic flight path gives a faster convergence of the bias estimates. Even if the IMU bias would be estimated perfectly, there would still be a slow error drift in position during total satellite outages, due to the accumulated sensor noise.

The atmospheric disturbance (ionospheric and tropospheric) affect the position estimates. The height component is significantly more affected than the two horizontal components.

8. Simulations of Jamming Suppression using Adaptive Beamforming

An adaptive beamforming antenna array can be combined with the integrated INS/GPS system.

The following simulations will be performed:

Different Numbers of Antenna Elements in the Array In Section 8.2 the results from a simulation with different number of elements in the array, is shown. A single LCMV beamformer is used. The equivalent carrier power-to-noise spectral density ration, $[C/N_0]$, are compared when using a beamforming antenna and an isotropic single element antenna.

Beamforming Strategies The performance of the system with a beamforming antenna array is tested under a realistic jamming scenario in Section 8.3. Three types of the chosen variance minimization beamforming algorithm are compared.

Robustness of Beamforming Algorithm Small errors in the *a priori* information to the beamforming algorithm will degrade the performance. In Section 8.4 it is investigated how large this degradation is.

The simulations are summarized in Section 8.5.

In the simulations the C/A-code will be used. The reason for this is that it is valid for all receivers. There are methods for using the P-coded and encrypted signal even if the encryption code is unknown, but this significantly decreases the SNR [9]. Therefore, in a situation with strong interfering signals, this will probably not be a useful method. In such a case, the user has to rely on the C/A-code measurement only. Actually, also most military receivers, with knowledge about the encrypted P-code, use the C/A-code signal when doing the first coarse acquisition of the signal.

Using only the C/A-code means that the performance is comparable with a civilian GPS receiver. For a military receiver using the P-code a better protection against jamming can be expected since the processing gain is 10 dB higher for the longer P-code.

8.1 General Settings

A rectangular array consisting of a number of isotropic antenna elements will be used in the simulations. The number of antenna elements in the array is varied. The L1 band will be used, since this is the frequency where the civilian C/A code is transmitted. Therefore the element spacing in the array is 0.09 m, which is slightly less than half a wavelength. Diagonal loading, described in Section 4.5.5, is used in all simulations. The load factor in equation (4.15) is set to $\alpha = 100$. This means that to the unloaded array correlation matrix, a diagonal matrix with elements 100 times the smallest eigenvalue of the array correlation matrix is added.

8.1.1 Settings for the INS When a dynamic scenario is simulated the navigation filter that integrates GPS and INS sensor data is used. The same inertial sensors parameters, corresponding to an Litton LN200 A1, as in Chapter 7 are used. The performances of the sensors are listed in Table 7.1 on page 59.

8.1.2 Settings for the GPS Receiver The GPS receiver is simulated by a rather simple model: a C/N_0 tracking threshold is set, and the tracking loops are assumed to lose lock if the equivalent carrier-to-noise spectral density $[C/N_0]$ is below this value. The quantity $[C/N_0]$, which describes the power of the GPS signal C, compared to the noise power spectral density N_0, is often used to describe the quality of the GPS signal instead of the SNR.

The critical part in the measurements is the acquisition of the GPS signal. Therefore another threshold, an acquisition threshold, is defined. This $[C/N_0]$ is required for the loops to first lock to the signal. The tracking threshold is set to 28.0 dBHz. It is taken from [3] and is valid for a L1 receiver with a pre-detection integration time of 20 ms, an Allan deviation of 10^{-11}, oscillator vibration sensitivity of 10^{-9} $\mathrm{s^2m^{-1}}$, and a maximum jerk stress of 100 $\mathrm{m/s^3}$. These assumptions are summarized in Table 8.1. For an explanation of these assumptions, the reader is referred to [3]. The acquisition of the GPS signal typically requires 5 dB higher $[C/N_0]$ than the tracking. Hence, the acquisition threshold is set to 33.0 dBHz.

The received power from the GPS satellites are all set to be -160 dBW, which is slightly less than the power that is guaranteed (-159.6 dBW, [3]). Note that the received power is often higher than the minimum guaranteed power. Losses in the receiver, e.g. from A/D conversion, is set to 4 dB.

The same settings for the generation of pseudorange error as for the simulations in Section 7.2 will be used. All properties are listed in Table 8.1. These settings are for a good quality C/A-code receiver. The model of the GPS receiver and how the pseudoranges are generated are described more in detail in Section 6.2.2. Settings for the beamformer are listed in Table 8.2.

Carrier frequency	L1 (1575.42 MHz)
Received satellite signal power	-160 dBW
Code used	C/A
Receiver losses	4 dB
Carrier loop tracking threshold	28.0 dB
Carrier loop acquisition threshold	33.0 dB
Carrier loop bandwidth	18 Hz
Pre-detection integration time (PIT)	20 ms
Allan deviation for oscillator	10^{-11}
Oscillator vibration sensitivity	10^{-10} $\mathrm{s^2}m^{-1}$
Maximum jerk stress	100 $\mathrm{m/s^3}$
Vibrations	0.5 $\mathrm{m^2s^{-3}}$

Table 8.1: Assumed parameters for GPS receiver in beamforming simulations.

Number of antenna elements:	varying
Array structure:	2–D square URA
Array element spacing:	0.09 m
Thermal noise power:	-140 dBW

Table 8.2: Assumed parameters for antenna array and beamforming algorithm.

8.1.3 Scenarios Three different scenarios will be used in the beamforming simulations. They are depicted in Figure 8.1.

8.2 Arrays with Different Numbers of Elements

The number of antenna elements in the array decides the degrees of freedom that can be used to form nulls. Since the element spacing is constant the number of elements will also determine the size of the array, the *array aperture*, which in turn determines the width of the main beam. The objective of this simulation is to show how the number of antenna elements affects the performance. This is tested for the scenarios described in Figures 8.1(a) and 8.1(b). The difference between the two scenarios is that there are 10 SNOIs in (a) and 15 SNOIs in (b).

A minimum variance beamforming algorithm (LCMV) has been used, with one constraint in the direction of the only SOI. The constraint is chosen correctly.

Comments The achieved SNRs are summarized in Table 8.3. The resulting beam patterns for the scenario with 10 SNOIs are shown in Figure 8.2. Arrays with a larger number of elements are able to form more narrow main beams.

Number of elements	SNR	
	10 jammers	15 jammers
1×1	-170.0 dB	-171.8 dB
2×2	-165.3 dB	-164.6 dB
3×3	-161.4 dB	-160.9 dB
4×4	-157.8 dB	-156.8 dB
5×5	-154.5 dB	-153.6 dB
6×6	-149.0 dB	-149.8 dB

Table 8.3: SNR for different number of elements. The scenarios in Figures 8.1(a) and 8.1(a) are used.

8.3 Comparing Beamforming Strategies

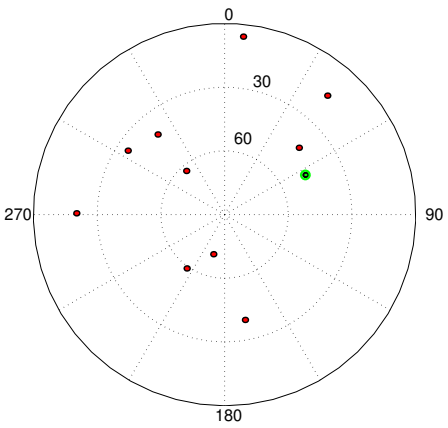
One of the beamforming algorithms discussed in Section 4.3, the power minimization algorithm, will be used with a (4×4) element URA. Three different strategies for setting the gain constraints will be compared. These three strategies are described in section 4.4.

Single LCMV Beamformer means that the LCMV power minimization algorithm is used with one single beamformer. The single beamformer must be able to form high antenna gains toward all satellites simultaneously and at the same time use the remaining degrees of freedom to suppress interfering signals.

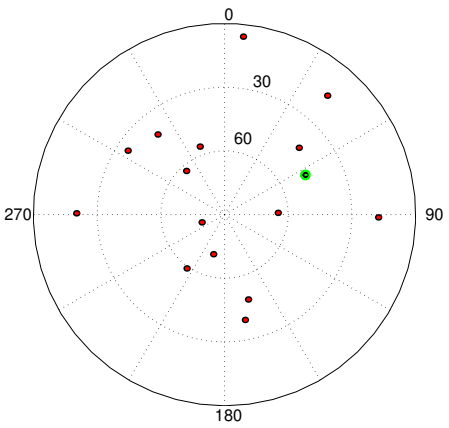
Multiple LCMV Beamformers means that the LCMV power minimization algorithm is used with several beamformers running in parallel. Each beamformer is optimized to direct its maximum gain towards one certain satellite.

Unconstrained Power Minimization This method uses a single beamformer which tries to minimize the output power. Hence, it works without any *a priori* information about the directions to the SOIs.

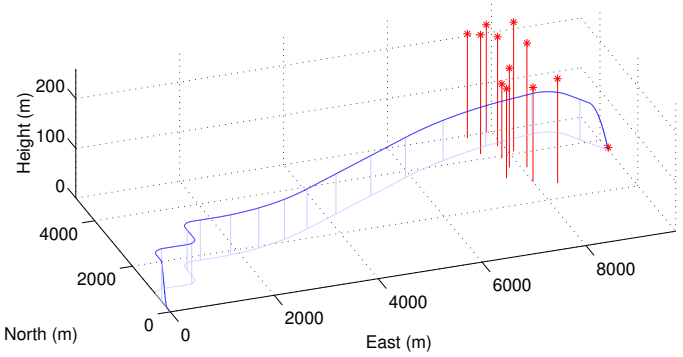
The directions to the satellites, that are *a priori* information to the first two types of beamformers, are calculated from the estimates given by the navigation filter. Estimates for the position and orientation of the vehicle are needed. In addition, the positions for the satellites are necessary. Equations (6.31) and (6.32) are used to calculate estimates of the azimuth ϕ and elevation θ, respectively.



(a) Scenario a. There are 10 SNOIs present. The only SOI is at direction $\phi = 64°$, $\theta = 48°$. The incoming power from each SNOI is 1 W (0 dBW). The incoming power from the SOI is -160 dBW. Hence, the total SNR with an isotropic antenna is -170 dB.

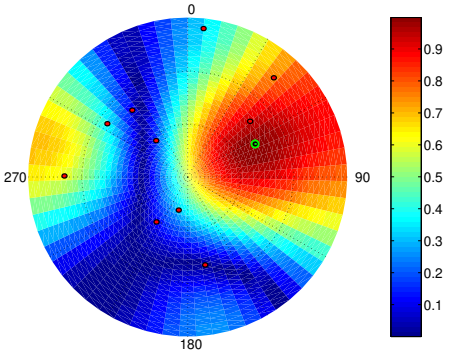


(b) Scenario b. This scenario is identical to scenario a, except that there are 15 SNOIs present. The total SNR with an isotropic antenna is now -171.8 dB.

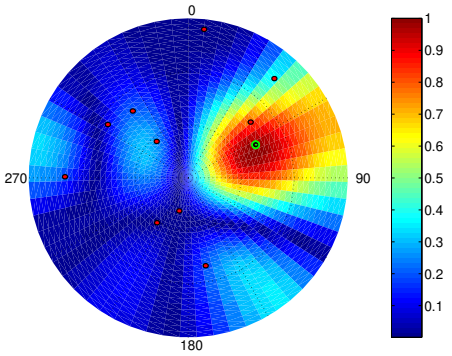


(c) Scenario c. In this scenario a flying vehicle enters an area with several jammers. A 10 W noise jammer is placed at the target (to the right) and 11 airborne low-power (100 mW) noise jammers are on the planned route at heights between 170 m and 260 m. The planned route starts on the ground (at $(0, 0)$), accelerates to about 100 m/s and then rises to 100 m above the ground. The received power from each satellite is -160 dBW. Free space propagation is assumed for the interfering signals. The projection of the flight path is projected on the ground and vertical lines are drawn every 10 seconds.

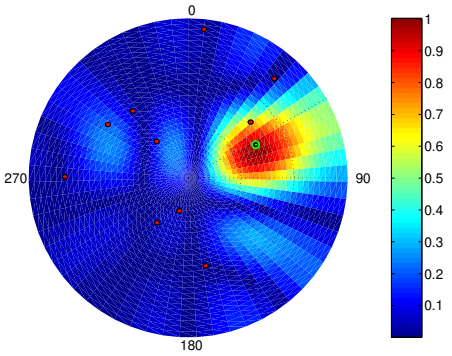
Figure 8.1: Scenarios used in beamforming simulations. In *Scenario a* and *Scenario b* all directions to signals are static. In *Scenario c* the directions are dynamic due to the movement of the vehicle and also the movement of the satellites.



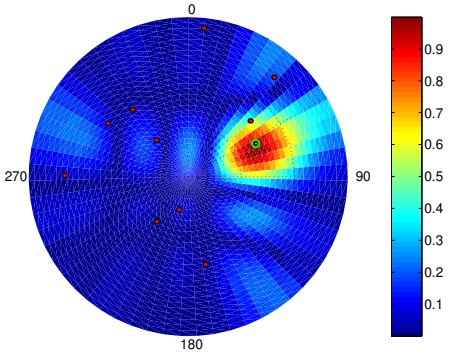
(a) 2×2 elements



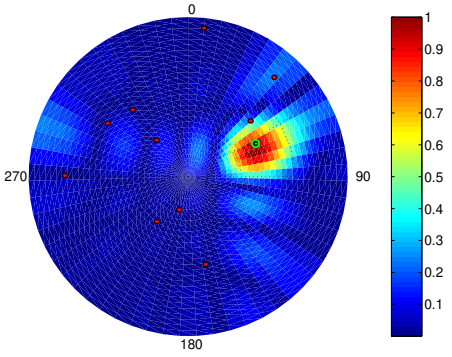
(b) 3×3 elements



(c) 4×4 elements



(d) 5×5 elements



(e) 6×6 elements

Figure 8.2: Gain patterns for different numbers of elements in the array. The scenario in Figure 8.1(a) is used. The only SOI is at direction $\phi = 64°$, $\theta = 48°$. The received power is the same from all ten interferers, 1 W (0 dBW) from each one. The received SOI power is -160 dBW.

The scenario depicted and described in Figure 8.1(c) will be tested. A flying vehicle navigating towards a certain point with a tight integration of an INS and a GPS enters an area with several jammers. Free space propagation is assumed for the interfering signals, i.e. equation (3.35) is used to determine the received power from the interfering sources. A (4×4) array is used. The results are compare with a single isotropic antenna.

Equation (3.37) is used to calculate the unjammed C/N_0. The thermal noise reference temperature T_0 is set to 300 K and the losses $L_s = 4$ dB. Since the received satellite power is assumed to be -160 dBW, the resulting unjammed C/N_0 is 39.8 dB. The beamforming antenna suppresses the interfering signals so that the jamming-to-signal ratio J/S is decreased. The $[C/N_0]$ after the beamforming is calculated using equation (3.39).

Beamforming method	Position error (m)	
	At end point	**RMS**
Isotropic antenna (no beamforming)	412 m	108 m
Single LCMV beamformer	315 m	67.5 m
Multiple LCMV beamformers	0.25 m	0.96 m
Unconstrained power minimization	1.40 m	1.01 m

Table 8.4: Position estimation errors using different beamformers. The scenario in Figure 8.1(c) is used. The numbers are averaged over 5 simulations

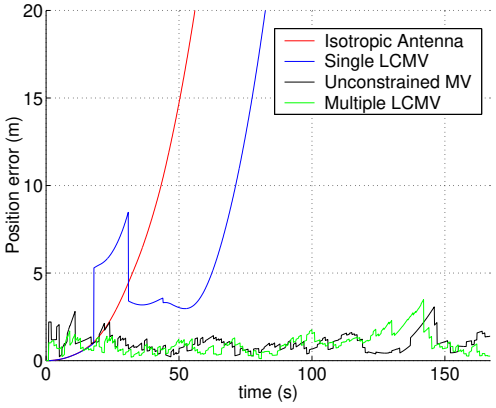
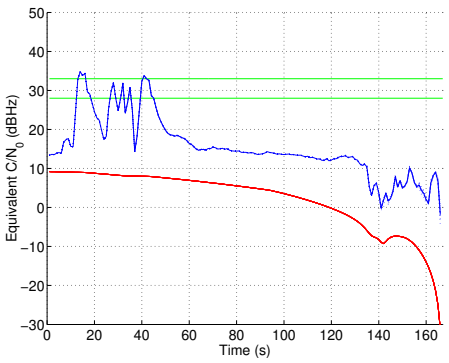


Figure 8.3: Position errors for the three beamforming methods compared. The position error when using a single isotropic antenna are also presented as a comparison.

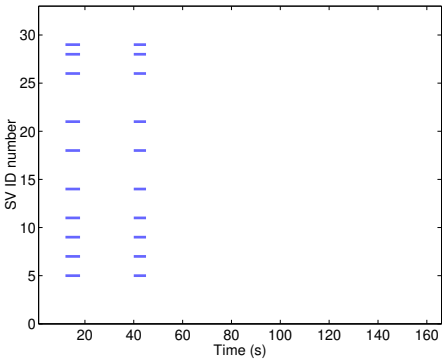
Results and Comments The resulting $[C/N_0]$ with the different adaptive beamforming algorithms are compared in Figure 8.4. The assumed SNR thresholds have the effect that not all visible satellites are possible to use all the time. The available satellites for the different algorithms are shown in Figure 8.4.

The resulting position errors are presented in Table 8.4 and Figure 8.3. The numbers in Table 8.4 are averaged over 5 simulations.

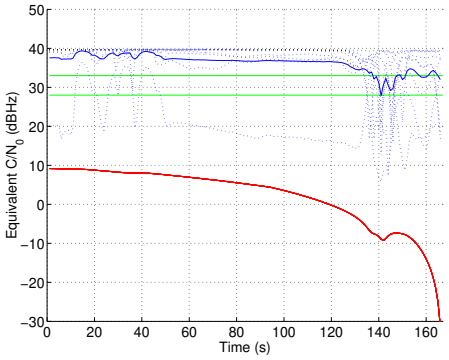
The reason for the bad performance of the single beamformer approach is that too many degrees of freedom is spent on setting gain constraints in the directions to all SOIs, and therefore not enough nulls can be formed to suppress jammers. The more constraints that are set in different directions, the more the gain pattern starts looking like a pattern from an isotropic antenna, since the gain constraints are spread over



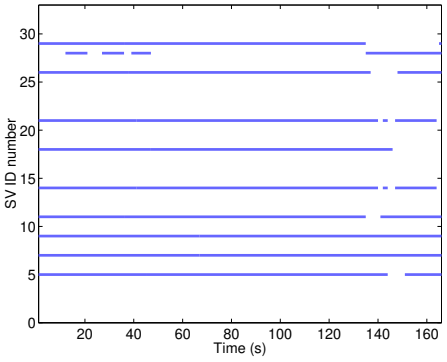
(a) Single LCMV beamformer. The average C/N_0 gain is 13.2 dB.



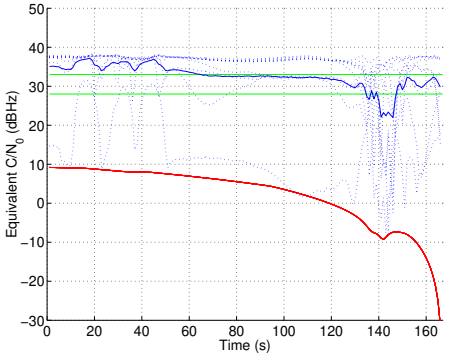
(b) Available satellites, single LCMV beamformer.



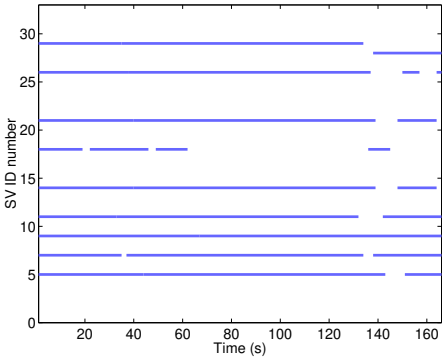
(c) Multiple LCMV beamformers. The average C/N_0 gain is 34.2 dB.



(d) Available satellites, multiple LCMV beamformers.



(e) Unconstrained power minimization. The average C/N_0 gain is 30.4 dB.



(f) Available satellites, unconstrained power minimization.

Figure 8.4: Comparing LCMV with different constraints. In (a), (c) and (e), the $[C/N_0]$ levels for each satellite when using an adaptive beamforming antenna is drawn with dotted lines. The mean $[C/N_0]$ for all visible satellites are drawn with a solid line (blue). The bottom solid lines (red) show the received $[C/N_0]$ if not using any beamforming, i.e. an isotropic antenna. A tracking threshold is drawn at 28.0 dBHz and an acquisition threshold at 33.0 dBHz (green). A satellite is available if it is visible and it can be tracked by the receiver.

almost the whole half-sphere. This has the effect that if only one interfering signal is close to any of the satellite's directions, this will increase the noise level not only for the signal from that satellite, but for all satellites. The $[C/N_0]$ will be equal for all satellites since the same received powers are assumed and the same gain is constrained in the directions of the satellites.

With multiple beamformers, one for each satellite, only one degree of freedom is spent on setting the gain constraint. Also, since there is one beamformer for each constraint, an interfering signal that is close to a satellite's direction will only increase the noise level for that satellite. The satellite availability is therefore significantly better compared to the single LCMV simulation.

The unconstrained power minimization beamformer also has a good satellite availability, resulting in small navigation errors. Having in mind that the calculation complexity is much lower for the unconstrained power minimization compared to multiple beamformers, the result is quite impressive.

8.4 Robustness for LCMV

The gain constraints that are set in the LCMV algorithm are the estimated directions to the satellites. Small errors in the estimated attitude of the vehicle means that the gain constraints in the direction to the satellites will be slightly wrong and, hence, the main lobes for all calculated beam patterns will be slightly misdirected. In this section it is investigated how attitude errors affect the resulting SNR.

The SNR is calculated for different errors in the estimated direction to the SOI. Errors in the estimated position also give a small error to the *a priori* information used by the beamforming algorithm. However, a position error must be unreasonably large to give any effect in the directions, because of the large distance to the satellites. Therefore, only attitude errors are tested.

Scenario a, with 10 SNOIs depicted and described in Figure 8.1(a) will be used. A constrained power minimization algorithm is used. Since only one SOI is present the previously described methods with single or multiple beamformers are equivalent.

The errors are introduced in the azimuth and elevation that describe the direction to the SOI. The introduced angle errors are Gaussian distributed noise with variance σ^2. The errors are introduced both in azimuth and elevation with $\sigma = \sigma_{azimuth} = \sigma_{elevation}$ and

$$\begin{aligned} \phi_{constraint} &= \phi_{true} + n_1 \\ \theta_{constraint} &= \theta_{true} + n_2, \end{aligned} \tag{8.1}$$

where ϕ_{true} and θ_{true} are the true directions to the satellites and $\phi_{constraint}$ and $\theta_{constraint}$ the directions used when setting the constraints. The two Gaussian noise processes n_1, n_2 both have variance σ^2. The degradation in SNR is tested for $\sigma = 0^\circ,\ 0.1^\circ, 1^\circ, 5^\circ, 10^\circ, 15^\circ, 20^\circ$. Three different arrays are tested, a (3×3) URA, a (4×4) URA and a (5×5) URA.

Results and Comments The SNR gain compared to a single element isotropic antenna are presented in Figure 8.5. The SNR gain for each array and introduced error σ is calculated as the average over 25 Monte Carlo simulations.

As expected, arrays with a larger number of elements give a higher SNR gain when the direction errors are small. This is due to the more narrow beams that can be formed. The degradation of SNR gain when errors are introduced is faster for larger arrays, also due to the more narrow beams that can be formed.

It is interesting to see that direction errors that can be expected from an integrated

INS/GPS (usually much less than 1°, see Section 7) hardly affect the SNR gain at all, not even the large (5×5) array.

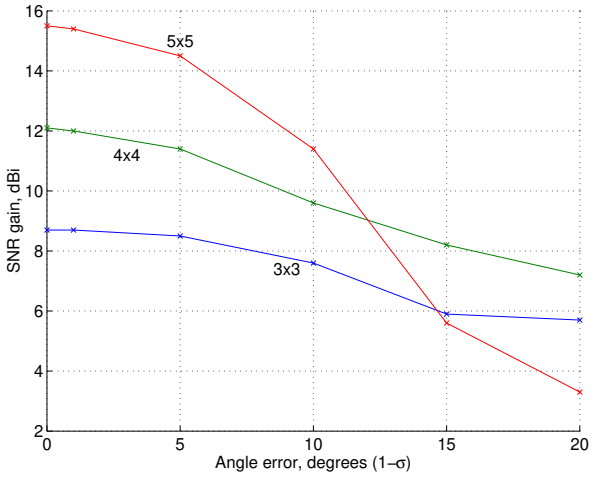


Figure 8.5: SNR gain compared to a single element isotropic antenna for different errors in the direction to the SOI. Three arrays are tested, a (3×3) URA, a (4×4) URA and a (5×5) URA. The scenario in Figure 8.1(a) is used.

8.5 Summary of Adaptive Beamforming Simulations

GPS receivers are extremely vulnerable to interfering signals. An adaptive antenna can increase the robustness against interference.

Using multiple beamformers, which all process the same sampled data gives the largest improvement of signal quality. With a single LCMV beamformer, it is an all-or-nothing situation. A good compromise in computation complexity and hardware cost is the unconstrained power minimization beamformer. It performs almost as good as the multiple LCMV beamformer in terms of navigation errors, but with much lower demands on hardware and computation capacity.

Two of the beamforming algorithms tested, single and multiple LCMV beamforming, require that the directions to the GPS satellites are known. It is shown that the the LCMV algorithm is robust in terms of perturbation in the constrained direction. The third beamforming algorithm, unconstrained power minimization, does not require any information about the directions to the satellite. Hence, it is insensitive to satellite direction errors.

9. Conclusions

Aiding an inertial navigation system with a GPS gives better navigation performance since the growing errors from the INS are bounded. A tight integration of INS and GPS has been implemented and is compared to a loose integration. With the tight integration, measurements from the GPS satellites can be passed to the filter also when less than four satellites are available.

The estimation of systematic errors and imperfections in the inertial sensors, e.g. gyro and accelerometer bias, is a big issue. These errors can be estimated by the navigation filter and be compensate for. A good estimate of the biases increases the navigation performance also during periods when no GPS aiding is available.

Using an adaptive beamforming antenna with a minimum variance algorithm, the equivalent carrier power-to-noise spectral density, $[C/N_0]$, can be significatively increased.

Three different types of a minimum variance algorithm have been implemented. Simulations have shown that with the most effective type, multiple LCMV beamformers, the $[C/N_0]$, can be increased with up to 40 dB on average, in a difficult signal environment. The unconstrained power minimization performs almost as good as multiple LCMV, but with much lower computation capacity needed.

The LCMV algorithm is in simulations shown to be robust to errors in the direction to the satellites.

9.1 Future Work

Several improvements can be made to get more realistic simulations:

More advanced GPS modelling With an improved GPS model which also includes the tracking loops, the simple $[C/N_0]$ thresholds for satellite availability could also be replaced with a more realistic criteria. If the loops are aided by the INS the jamming susceptibility can be decreased. With real tracking loops it can also be tested how fast the beamforming weights must be updated to avoid phase slips.

In [27] a software GPS receiver that works in real-time is described. This implementation is also prepared to sample and handle real data.

Signal propagation model Today, free space propagation is assumed for the interfering signals and a constant received power from the satellites. A more detailed signal propagation model makes the simulations more realistic.

Integrity control The integrity of the GPS measurements can be monitored by comparing INS and GPS data. It is important to be able to tell system errors from performance degradation caused by jamming or spoofing.

Control unit The navigation filter could be integrated with a control unit that uses the estimated position to navigate the vehicle to a certain point. An interesting application would be to recalculate the planned flight depending on the constellation of hostile jammers.

Better user interface With a user interface and a connection to map and terrain data, it could be possible to define flight paths, place jammers directly on the map, etc.

A. Notations

General Notations

$\mathbf{a}$	Vector (bold face, lower case)
$\mathbf{A}$	Matrix (bold face, upper case)
a	Scalar
$\hat{(\cdot)}$	Estimate
$\dot{(\cdot)}$	Time derivative
$\ddot{(\cdot)}$	Second time derivative
$(\cdot)^*$	Complex conjugate
$(\cdot)^T$	Transpose
$(\cdot)^H$	Hermitian transpose
$\tilde{(\cdot)}$	Nominal solution
$\delta(\cdot)$	"Small" error
$[(\cdot)\times]$	Skew-symmetric matrix
$(\cdot)_{k\|k-n}$	Prediction of $(\cdot)_k$ given $(\cdot)_{k-n}$
$E[\cdot]$	Expected value for ensembles
$\varepsilon[\cdot]$	Expected value for time series
$\|\cdot\|$	Vector norm

Constants

c	Speed of light	299 792 458 m/s
f_1	Carrier frequency, L1 band	1575.42 MHz
f_2	Carrier frequency, L2 band	1227.60 MHz
ω_{ie}	Earth rotation rate	$7.292115 \cdot 10^{-5}$ rad/s
M	Mass of Earth	$5.972 \cdot 10^{24}$ kg
k	Gravitational constant	$6.67 \cdot 10^{-11}$ $\mathrm{m}^3/(\mathrm{kg\,s}^2)$

Notation of Variables

t	Time-continuous index
k	Time-discrete index
$\mathbf{x}$	Position
$\mathbf{v}$	Velocity
$\mathbf{\Psi}$	Orientation
$\mathbf{d}$	Constant gyro bias
$\mathbf{b}$	Constant accelerometer bias
$\boldsymbol{\varepsilon}$	Error state vector in Kalman filter
$c\,\delta t$	GPS receiver clock offset
$c\,\delta t_s$	Satellite clock offset in satellite s
$\mathbf{w}$	Weight vector in antenna array
$\mathbf{e}$	Unit vector, $\|\mathbf{e}\| = 1$
$\mathbf{s}$	Steering vector
σ^2	Variance
ρ	Code pseudorange
φ	Phase pseudorange

v	Gaussian noise in pseudorange measurement
r	Geometric distance
ϕ	Azimuth angle
θ	Elevation angle
$\mathbf{z}$	Arbitrary state vector in Kalman filter
$\mathbf{0}_n$	$(n \times n)$ matrix with zeroes
$\mathbf{I}_n$	$(n \times n)$ identity matrix
α	Load factor in diagonal loading of a correlation matrix
$\boldsymbol{\mathcal{O}}$	Observability matrix
$\mathbf{\Phi}$	Discretized system matrix in Kalman filter
$\mathbf{F}$	System matrix in Kalman filter
$\mathbf{G}$	System noise matrix in Kalman filter
$\mathbf{H}$	Measurement matrix in Kalman filter
C/N_0	Unjammed carrier power-to-noise spectral density
$[C/N_0]$	Equivalent carrier power-to-noise spectral density

B. Abbreviations

BPSK	Binary Phase Shift Keying
C/A	Coarse/Acquisition
CDMA	Code Division Multiple Access
DCM	Direction Cosine Matrix
DGPS	Differential GPS
DOA	Direction of Arrival
DOP	Dilution of Precision
ECEF	Earth-Centered Earth-Fixed
ECI	Earth-Centered Inertial
EKF	Extended Kalman Filter
ENU	East-North-Up
FLL	Frequency Locked Loop
FOI	Swedish Defence Research Agency
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
IF	Intermediate Frequency
INS	Inertial Navigating System
IMU	Inertial Measurement Unit
LCMV	Linear Constraint Minimum Variance
LKF	Linearized Kalman Filter
MMSE	Minimum Mean Square Error
NCO	Numerical Controlled Oscillator
NED	North-East-Down
PLL	Phase Locked Loop
PPS	Precise Positioning System
PRN	Pseudo-Random Noise
RF	Radio Frequency
RMS	Root Mean Square
SA	Selective Availability
SNR	Signal-to-Noise Ratio
SNOI	Signal Not of Interest
SOI	Signal of Interest
SPS	Standard Positioning System
SV	Space Vehicle
SVID	Space Vehicle Identity
TOA	Time of Arrival
UERE	User Equivalent Range Error
ULA	Uniform Linear Array
UCA	Uniform Circular Array
URA	Uniform Rectangular Array
UTC	Coordinated Universal Time

References

[1] S-L. Wirkander B. Boberg. Integrating GPS and INS: Comparing the kalman estimator and the particle estimator. Technical report, Swedish Defence Research Agency, June 2002.

[2] Lars Pettersson. WBBF ett simuleringsverktyg för bredbandig adaptiv lobformning. Technical report, Swedish Defence Reseach Agency (FOI), June 2002.

[3] E. D. Kaplan. *Understanding GPS, Principles and Application.* Artech House Publishers, 1996.

[4] C. Jekeli. *Inertial Navigation System with Geodetic Applications.* de Gruyter, 2001.

[5] Robert Grover Brown Patrick Y. C. Hwang. *Introduction to Random Signals and Applied Kalman Filtering.* John Wiley & Sons, Inc, third edition, 1997.

[6] Bjørnar Vik. *Nonlinear Design and Analysis of Integrated GPS and Inertial Navigation Systems.* PhD thesis, NTNU Trondheim, Institutt for teknisk kybernetikk, 2000.

[7] Jay A. Farrell Matthew Barth. *The Global Positioning System & Inertial Navigation.* McGraw-Hill, 1999.

[8] Martin Davidson. Simulation of interfering signal in direct sequence systems concentrating towards GPS. Master's thesis, Umeå University, Department of Computing Science, June 1996.

[9] B. Hofmann-Wellenhof et al. *GPS, Theory and Practice.* SpringerWienNewYork, fourth edition, 1997.

[10] M. Horemuž L.E. Sjöberg. Rapid GPS Ambiguity Resolution For Short and Long Baselines. *J. Geod.*, Vol. 76(no. 6–7):381–391, 2002.

[11] Anders Alm. GPS och störskyddsantenner. Technical report, FOA, Avdelningen för sensorteknik, Linköping, January 1995.

[12] D. J. Jwo. Optimisation and sensitivity analysis of GPS receiver tracking loops in dynamic environments. *IEEE Proc.-Radar, Sonar Navig.*, Vol. 148(no. 4), August 2001.

[13] Lars Ahlin Jens Zander. *Principles of Wireless Communications.* Studentlitteratur, Lund, second edition, 2000.

[14] Z. Fu A. Hornbostel J. Hammesfahr A. Konovaltsev. Suppression of multipath and jamming signals by digital beamforming for GNSS/Galileo application. Technical report, Deutsches Zentrum fuer Luft- und Raumfahrt. e, V (DLR).

[15] Simon Haykin. *Communication System Engineering.* Prentice Hall, New Jersey, 1994.

[16] Michael D. Zoltowski Anton S. Gecan. Advanced adaptive null steering concepts for GPS. *IEEE*, 0-7803-2489(7/95):1214–1218, 1995.

[17] J.C. Liberti T.S. Rappaport. *Smart antennas for wireless communications: IS-95 and third generation CDMA applications*. Prentice, 1999.

[18] Fredrik L. T. Eklöf Lars E. Pettersson Jouni G. J. Rantakokko. On broadband adaptive beamforming for tactical radio system. In *IEEE Military Communication Conference (MILCOM)*, Washington D.C., October 2001.

[19] Barry D. Van Veen Kevin M. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, April 1998.

[20] Mårten Lindgren. GPS och adaptiv antenn. Technical report, FMV, KC Sensor & Tele, May 2002.

[21] B. D. Carlson. Covariance matrix estimation error and diagonal loading in adaptive arrays. *IEEE Transaction on Aerospace and Electronic System*, 24(4), July 1988.

[22] P. S. Maybeck. *Stochastic Models, Estimation, and Control*, volume 141-1. Mathematics in Science and Engineering, 1994.

[23] GPSoft. *Satellite Navigation Toolbox 2.0, User's Guide*, 1999.

[24] F. Gustafsson. *Adaptive Filtering and Change Detection*. Wiley & Sons, 2000.

[25] T. Glad L. Ljung. *Reglerteori, Flervariabla och olinjära metoder*. Studentlitteratur, Lund, 1997.

[26] Swedish Defence Material Administration (FMV). Skrovfasta sensorer fpl 37 800. May 1979.

[27] P.-L. Normark C. Stahlberg. Real-time global navigation satellite system (GNSS) software receiver. Technical report, Lulea University of Technology.